

FIG. 3A-1

350 JD13 PAATLE EMMTACQGVG GPGHKARVLA EAMSQVTNS- ATIMMQRGNF RNQRKIVKCF G S A T A KGP- RI C	430 440 450 460 470 480 EGHQM KDCTERQANF LGKIWPSYKG RPGNFLQSRP EPTAPPFLQS RPEPTAPPEE L R H	S
380 EAMSQVTNS- AT A T A A V T	460 RPGNFLQSRP EP	6-86 PI
370 GPGHKARVLA S S	450 LGKIWPSYKG R H	
360 EMMTACQGVG	440 KDCTERQANF	510 TSLRS LFGNDPSSQ A K QL
350 KALGPAATLE G	430 WKCGKEGHQM R R R	
330 340 NWMTETLLVQ NANPDCKTIL KALGP	420 NCRAPRKKGC	500 PSOKQEPIDK ELYPL QK QK
330 NWMTETLLVQ	410 420 NCGKEGHIAR NCRAPRKKGC WKCGK K L K	490 SFRSGVETTT I F E K GF E IK- GF E I -
LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI

Central Region: Q

80 GLHTGERDWH Q K	160 AALITPKKIK T T A TR T A Q	
70 DARLVITTYW K VR E K	150 KVGSLQYLAL	
60 ISSEVHIPLG V V	140 PRCEYQAGHN D	FIG. 38-1
30 40 50 60 70 80 VKHHMY VSGKARGWFY RHHYESPHPR ISSEVHIPLG DARLVITTYW GLHTGERDWH I K K V K V E K V K V E K V E E K O K E	90 100 110 120 130 140 150 160 LGGGVSIEWR KKRYSTQVDP ELADQLIHLY YFDCFSDSAI RKALLGHIVS PRCEYQAGHN KVGSLQYLAL AALITPKKIK KN I YR H Q L D R L D F Q I D R T A TR	E E
40 VSGKARGWFY I K K K KN K NR	120 YFDCFSDSAI E	НЭ
30 WKSLVKHHMY H	110 ELADOLIHLY G D G	190 KGHRGSHTMN .R
20 WQVDRMRIRT K	100 KKRYSTQVDP 'K Q L	180 EDRWNKPQKT Q Q
10 20 MENRWQVMIV WQVDRMRIRT WKSL H	90 LGGGVSIEWR H	170 180 190 PPLPSVTKLT EDRWNKPQKT KGHRGSHTMN GH K R R Q R
LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI

FIG. 3B-1

 \propto

80 IFRIGCRHSR Q Q	, ·	E K0
30 40 50 60 70 80 80 EELKNEA VRHFPRIWLH GLGQHIYETY GDTWAGVEAI IRILQQLLFI HFRIGCRHSR S S Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q		30 40 50 60 70 TTCYCKKC CFHCQVCFTT KALGISYGRK KRRQRRRPPQ GSQTHQVSLS KQ NN YARG NK YMIG GADPPE NK HYPLN G
6DTWAGVEAI F		60 KRRGRRRPPQ A G
50 GLGOHIYETY S S S		50 KALGISYGRK G G G
40 VRHFPRIWLH P	·	40 CFHCQVCFTT YAR Y M I Y P LN
30 ELLEELKNEA R 0 S		30 ACTTCYCKKC NN P NK P NK H
10 20 MEGAPEDGGP GREPHNEWTL Y A Y A	-NGASRS R - S - S - S	MEPVDPRLEP WKHPGSQPKT ACT N N R P N D N N R P N D N N R P N
10 MEQAPEDQGP A	90 IGVTQQRRAR -NGASRS II R - S IIR - S S (tat)	10 MEPVDPRLEP N D N
LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI

ŗ.

FIG. 3B-2

POL

IKIGGQLKEA LLDTGADDTV LEEMSLPGRW KPKMIGGIGG FIKVRQYDQI LIEICGHKAI GTVLVGPTPV NIIGRNLLTQ R N K VRV VRV FFREDLAFLO GKARÈFSSEO TRANSPTFSS EQTRANSPTR RELOVWGRDN NSLSEAGADR OGTVSFNFPO ITLWORPLVT IGCTLNFPIS PIETVPVKLK PGMDGPKVKQ WPLTEEKIKA LVEICTEMEK EGKISKIGPE NPYNTPVFAI KKKDSTKWRK -VDFRELNKR TQDFWEVQLG IPHPAGLKKK KSVTVLDVGD AYFSVPLDED FRKYTAFTIP SINNETPGIR YQYNVLPQGW K لنالنا \propto р 6 L PK BRU 2 MAL ELI LACA LACA V

	•	•
400 EPPFLWMGYE	480 AENREILKEP	560 GKTPKFKLPI · I R R R
390 LTTPDKKHQK F F F R	470 PLTEEAELEL A	550 KITTESIVIW VS AQ R S
350 350 400 IVIYQY MDDLYVGSDL EIGQHRTKIE ELRQHLLRWG LTTPDKKHQK EPPFLWMGYE K F F R	410 420 430 440 450 460 470 480 LHPDKWTVQP IVLPEKDSWT VNDIQKLVGK LNWASQIYPG IKVRQLCKLL RGTKALTEVI PLTEEAELEL AENREILKEP M S K D E S K E E N ER	540 DVKQLTEAVQ
370 EIGQHRTKIE	450 IKVRQLCKLL K K	530 YARTRGAHTN M IKS M
360 MDDLYVGSDL	. 440 LNWASQIYPG	520 EPFKNLKTGK '
40 350 RK QNPDIVIYQY TKEM	430 VNDIQKLVGK N ER	510 QGQWTYQIYQ H
330 340 KGSPAIFQSS MTKILEPFRK QNPD T K E	420 IVLPEKDSWT M Q D E K	500 DLIAEIQKQG (
330 KGSPAIFQSS	410 LHPDKWTVQP S	490 500 510 520 530 540 550 560 VHGVYYDPSK DLIAEIQKQG QGQWTYQIYQ EPFKNLKTGK YARTRGAHTN DVKQLTEAVQ KITTESIVIW GKTPKFKLPI M OY IKS A R S R R
LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI

FIG. 3C-2

640 QKVVTLTDTT SIA SPE	720 GNEQVDKLVS	806 DCTHLEGKVI I	880 NPQSQGVVES
570 580 590 600 610 620 630 640 QKETWETWWT EYWQATWIPE WEFVNTPPLV KLWYQLEKEP IVGAETFYVD GAASRETKLG KAGYVTNRGR QKVVTLTDTT A M K D SIA I N K D SIA	670 680 690 700 710 720 DSQYAL GIIQAQPDKS ESELVNQIIE QIIKKEKVYL AWVPAHKGIG GNEQVDKLVS I Q D S	750 760 800 NWRAMA SDFNLPPVVA KEIVASCDKC QLKGEAMHGQ VDCSPGIWQL DCTHLEGKVI I	830 840 850 860 870 880 AYFLLK LAGRWPVKTI HTDNGSNFTS TTVKAACWWA GIKQEFGIPY NPQSQGVVES I VV AA AA N
620 GAASRETKLG N N N	700 QIIKKEKVYL Q D	780 QLKGEAMHGQ	860 TTVKAACWWA AA AA
610 IVGAETFYVD I	690 ESELVNQIIE I	770 KEIVASCDKC	850 HTDNGSNFTS
600 KLWYQLEKEP T	680 6110AQPDKS	760 SDFNLPPVVA I	840 LAGRWPVKTI VV VV
590 WEFVNTPPLV	670 NIVTDSQYAL	750 KYHSNWRAMA N	830 GQETAYFLLK I
580 EYWQATWIPE	650 660 NQKTELQAIH LALQDSGLEV NIVT S	740 GIDKAQDEHE E E	820 IEAEVIPAET
570 QKETWETWWT A M A	650 NQKTELQAIH N	730 740 AGIRKVLFLD GIDKAQDEHE S Q E	810 820 LVAVHVASGY IEAEVIPAET GQET
LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI

FIG. 3D-1

096	VYYRDSRDP	¥z ·			
950	MNKELKKIIG QVRDQAEHLK TAVQMAVFIH NFKRKGGIGG YSAGERIVDI IATDIQTKEL QKQITKIQNF RVYYRDSRDP N	Ι			
940	IATDIQTKEL			RQDED	9
930	YSAGERIVDI	W III	1010	KVVPRR KAKIIRDYGK QMAGDDCVAS RQDED	9 9
920	NFKRKGG1GG	RR	1000	KAKIIRDYGK	>
910	TAVQMAVFIH		066	NSDIKVVPRR	¥
006	QVRDQAEHLK	ш	086	KGEGAVVIQD	
890	MNKELKKI 1G N		970	LWKGPAKLLW KGEGAVVIQD NSDI	
	BRU 2	LAV MAL LAV ELI			LAV MAL LAV ELI

FIG. 3D-2

EN

	80 «ATHACVPT	160 SSSGEMMME- NWKE I RTNA LK I TTEEKG	240 YCAPAGFAI T	320 TRPNNNTRK G R A YQ Q
	30 40 50 80 70 80 80 IN SATEKLWVT VYYGVPVWKE ATTTLFCASD AKAYDTEVHN VWATHACVPT R SE A I		210 220 230 240YTLTS CNTSVITGAC PKVSFEPIPI HYCAPAGFAI TNYTN R IN RS R IN AS R IN A	270 280 300 310 320 IRPVV STOLLLNGSL AEEEVVIRSA NFTDNAKTII VOLNOSVEIN ČTRPNNNTRK K FIG. 3E-1
OMP	60 ATTTLFCASD ,	90 100 110 120 130 140 150 DPNPQEVVLV NVTENFNMWK NDMVEQMHED IISLWDQSLK PČVKLTPLČV SLKČTDL-CN ATNTNSSNTN C N Q TNACS IE E G N CTNACS IA E G N CTNACS	220 CNTSVITGAC	300 NFTDNAKTI:I N T N L N N
	50 VYYGVPVWKE	130 PČVKLTPLČV	210 YTLTS TNYTN R IN S R IN	290 AEEEVVIRSA IM E IM E
	CSATEKLWVT IA D ADN	120 IISLWDQSLK	190 200 FFYKLD IIPIDNDTTS L RN VV AS T T N LVQ DSDN L R V SST	270 280 2VV STOLLLNGSL 1 FIG. 3E-
SP	6TMLL - M - 1	110 NDMVEQMHED N N	190 KEYAFFYKLD N L RN OV L R	270 0CTHGIRPWV 1 K
	20 QHLWRWGWKW NW	100 NVTENFNMWK G	180 ISTSIRGKVQ KEYAFF T D I N L TPVGSD R - T VT VLKD K QV L	250 260 LKČNKKTFNG TGPČTNVSTV QČTHG RD K RD K EI K
	10 MRVKEKY K GTRRN REI ORN ARGIERNO	90 DPNPQEVVLV C IE E IA E	170 KGEIKNČSFN W	250 LKČNKKTFNG D K RD K
	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI

350 360 370 380 390 400 NMRQAHCNIS RAKWNATLKQ IASKLREQFG NNKT-IIFKQ SSGGDPEIVT HSFNGGGEFF DI K O V GSLL- K NS T R R CTLL- I K P	CSNNTEGSDT ITLPCRIKQF INMWQEVGKA MYAPPISGQI RCSSNITGLL LTRDGGNN RTEG K N I KT A V N L I NSSD T = S STGS I K VAGR- I ERN L I NSSD	510 520 530 TMP 550 560 SYKVYKI EPLGVAPTKA KRRVVQREKR AVGI-GALFL GFLGAAGSTM GARSMTLTVQ V L V M V L V M M V L V L M A L A L A L
380 IFKQ S: V N K P	460 SGQ1 RC C S A V N ERN L	540 TA ALFL GF M
NNKT-I - K - I	MYAPP I	AVG1-6 V I L- I L-
370 IASKLREGFG VK V V GSLL- V R GTLL-	450 INMWQEVGKA KT K VAGR-	530 KRRVVQREKR A E
360 RAKWNATLKQ Q N E ETE DK Q	440 ITLPCRIKQF I Q I	520 EPLGVAPTKA I
350 NMRQAHCNIS DI K DI R Y T N IIG	430 CSNNTEGSDT RTEG K N TES STGS	510 ELYKYKVYKI I R Q
340 AFVTIGK-IG W T RI LY T I-V SLY TKS-RS		500 GGDMRDNWRS I
SIRIQRGPGR AFVTIGK-IG Y Q LY T RI G HF Q LY T I-V RTP L Q SLY TKS-RS	410 420 YCNSTQLFNS TWFNSTWSTE T NRLN TSK Q NGARL- TSG NI A NNI	490 500 NNGSEIFRPG GGDMRDNWRS ELYKY T DT V SDN TL I
LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI

FIG. 3E-2

640 KSLE R D R N	720 VFA I	800 IVT AA A AV	
NASWSNK S	GLVGLRI I I	RLRDLLL	GLERILL F A S
630 KLIČTTAVPW H F	710 YIKIFIMIVG R IV	790 LRSLCLFSYH N	870 IRHIPRRIRQ L L L VLN
620 QLLGIWGCSG R M	700 NWFNITNWLW S SK S Q	780 NGSLALIWDD D F FS FS	860 IEVVQGAÇÜRA A R Y IĞ RFĞ II R
570 580 600 610 620 630 640 AROLLSGIVQ QQNNLLRAIE AQQHLLQLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KLICTTAVPW NASWSNKSLE M M M M M M M M M M M M M	690 LELDKWASLW	750 780 800 800 800 PGGP-DR PEGIEEEGGE RDRDRSIRLV NGSLALIWDD LRSLCLFSYH RLRDLLLIVT AA AA AA T OG G V L FS N I AV I AV I A	820 830 840 850 860 870 WEALKYWWNL LQYWSQELKN SAVSLLNATA IAVAEGTDRV IEVVQGAGRA IRHIPRRIRQ GLERILL DI L G I T G L F A F A B C C C C C C C C C C C C C C C C C C
60.0 WGIKQLQARI W	680 NQQEKNEQEL I K	. 760 PEGIEEEGGE D	840 SAVSLLNATA W I T S FD I
590 AQQHLLQLTV	670 LINSLIEESQ T YT L I YN	750 LPTPRGP-DR V P A -	830 LQYWSQELKN I G
580 QQNNLLRAIE	660 WDREINNYTS E D N EK S G	740 GYSPLSFQTH R L L	820 WEALKYWWNL S DI L
570 ARQLLSGIVQ	650 660 670 680 690 700 710 720 QIWNNMTWME WDREINNYTS LINSLIEESQ NQQEKNEQEL LELDKWASLW NWFNITNWLW YIKIFIMIVG GLVGLRIVFA S G I YN I K	730 740 VLSIVNRVRQ GYSPLSFQTH LPTF R V L L L A	810 RIVELLGRRG T I K
LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI

FIG. 3F-1

30 40 50 60 70 80 80 PV SNFAPT NAACAWLEAG EE-EEVGFPV D D AVSG D C AA SP N S PP E SD SD	110 120 130 140 150 160 GGLEGL IHSQRRQDIL DLWIYUTQGY FPDWQNYTPC PGVRYPLTFG WCYKLVPVEP W F F F HS D VW PK E V N I I F F F HS	
60 AITSSNTAAT AA SP N	140 FPDWQNYTPC	EYFKNC Y D Y D FY _
50 DLEKUG AVSQ D C	130 DLWIYUTQGY V N I	210 FHHVARELHP M LR R Q LE K M
40 ADGVGAASR- V ET V QD	120 IHSQRRQDIL W PK E W KK E	190 200 210 DPEREV LEWRFDSRLA FHHVARELHP EYFKNC A K V K S LR R Q Y D A K K S LR R Q Y D A K K S LR R Q Y D
RAE	110 LKEKGGLEGL D	190 HGMDDPEREV E A K E A Q
20 VGWPTVRERM G SAI KII I	100 YKAAVDLSHF L I G F E L	JZO J80 SKVEEANKGE NTSLLHPVSL HGMDI EE E NC I Q E SE DTE TN ICQ E
10 20 MGGKWSKSSV VGWPTVRERM F R M G SAI I KI I I AI I	90 100 TPQVPLRRHT YKAAVDLSHF LKEK(R R G F R	J70 DKVEEANKGE E E GE DTE
F LAV BRU LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI	LAV BRU ARV 2 LAV MAL LAV ELI

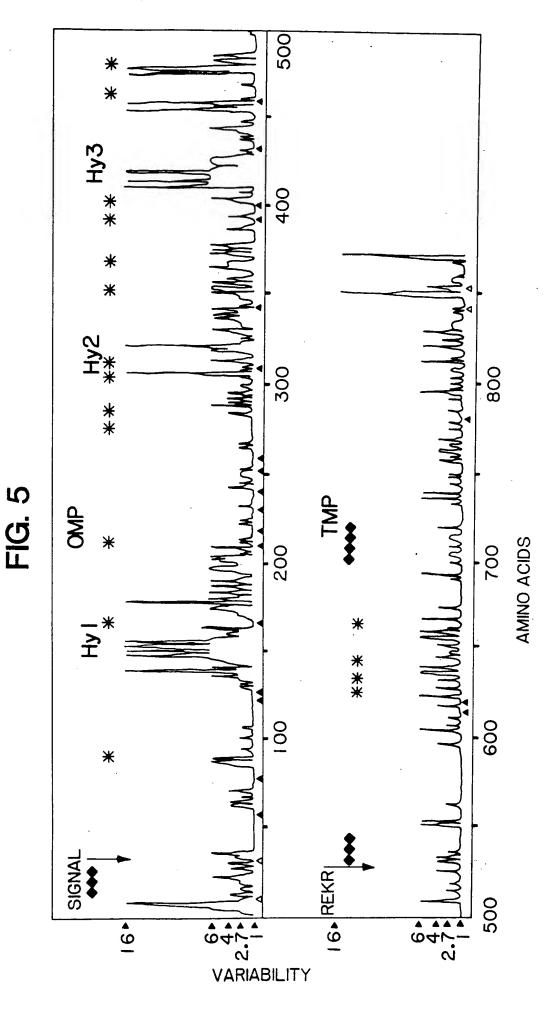
FIG. 3F-2

FIG. 4A

Ø	A LAVbru		0	۵	2			ENV	>		
'	VS.	5	Ş	•	7	T0.	TOTAL	OMP	Ψb	F	ТМР
	HTLV-3 512 USA 0/0	512 0.8 1015 0/0	0.8	0/0	1.3 856	856 570	1.4 507 1.6 349 5/0 0/0	507	9°1	349	-
·	ARV-2 USA	502 12/2	3.4	1003	1003 3.1	855 17/11	855 13.0 505 14.3 350 1	505 1	14.3	350	11.2
	LAVeli ZAIRE	500 9.8	9.8	1002 13/0	5.5	853 2 (853 20.7 504 25.3 349 13.8 22/14	504 25. 22/14	25.3	349	13.8
<u> </u>	LAVmal 505 2.0 1002 7.7 859 21.7 509 26.4 350 14.9 2AIRE 14.7 13.7 10 13.7 10 10.1	505	2.0	1002	7.7	859 3/11	21.7	509 3/10	26.4	350	14.9
m	LAVeli vs.										
	LAVmal	505 1/6	0.8	005	505 10.8 1002 8.4 859 19.8 509 23.6 350 14.3	59 5	9.8 5 8,	509 2.	3.6	50	4.3

FIG. 4B

A	A LAVbru	orf F	L		central region	l regi	on		
	VS.			Or.	orf Q	0	orf R	ဝ	orf S
	HTLV-3 USA	206 070	1.5	192 070	0		Ъп	80 0/0	80 2.5
	ARV-2 USA	210	2 210 12.6 192 10.0 97 9.4 81 15.0 USA 0/4 15.0	192 070	0.0	97	9.4	81	15.0
	LAVeli 206 19.4 192 10.4 96 1.1.5 80 27.5	206 171	19.4	192	10.4	96	7.5	80	27.5
	LAVmal 209 27.0 192 12.6 96 10.4 80 23.8	209	27.0	192	12.6	96	10.4	80	23.8
m	LAVeli vs.								
	LAVmal	209	209 22.5 192 12.0 96 3/6 0/0 12.0 0/0	192	12.0	96	96 6.3 80 11.3	80	5.



6AG

_

 K
 AAA
 GCA
 CAA
 GCA
 GCA D T GAC ACA LAV.BRU LAV.MAL LAV, EL I ARV 2

FIG. 6A-1

LAV.BRU

Q

480	E GAG	E 6A6	E GAG	E GA6
	GCC CCA CCA TIT CIT CAG AGC AGA CCA GAG CCA ACA GCC CCA CCA	E E E	A E GCA GAG	GCA GAG
	CCA	1	ı	1
	CCA	ı	t	ı
	A GC(1	ť	1
	A AC	ı	1	ı
024	9 C CC	1	1	1
	A GA	1	ŧ	1
	A CC	1	ı	1
	C AG	ı	1	1
	16 AG	÷ 1		ı
	TT CA	1	ı	ı
] [3 [4]	1	1	1
	P CA		- Y	- Y
	P CA C	A P P GCC CCA	A P P	A P P GCC CCA CCA
	A CC C	ر ر ک	CC C(۸ در در
	P CA A	P CA A	P CA A	P CA A
	E 3AG (E 3AG (E iAG C	E AG C
	CCA (P CCA	P SCA	P CA G
	R AGA	R AGA	R AGA (R 4GA (
	S AGC	S AGC	S AGC	S AGC /
	CAG	Q CAG	0 CAG	CAA
	CTT	L CTT	L CT	CTC
	1	ARV'2 G N F L Q S R P E P T GGG AAT TTT CTT CAG AGC AGA CCA GAG CCA ACA	LAV.MAL Genat Tic cit cae agc aga cca gae cca aca	LAV.ELI G N F L Q S R P E P T GGG AAC TTT CTC CAA AGC AGA CCA GAG CCA ACA
	N AAT	2 N AAT	LAV.MAL G GGG AAT	LAV.ELI GGG AAC
	999	ARV 2 G N	LAV 666	LAV.

FIG. 6A-2

30	A GCA	A GCA	ACA	T GCA		D GAC	D GAC	D GAT	D GAC
	CCA	CCA	CCA	CCA		ı	ı	CAA	1
	E GAG	E GAG	^ه 223	PAAT		* 1	ı	STCT	1
	A GCT	GCT (T ACT	TACT		1	1	V S GTA TCT	1
	R CGA	CGA	R CGA	R AGA		1	1	SCA	1
	ı	CCA	1.	ı		1	1	D	ı
	1	E GAG	1	ŧ	04	R CGA	R CGA	SAA	R CGA
	1	AGA CGA GCT	1	1		S TCT	·A TCT	ST	S TCT
	1	CGA	1	l		A GCA	V GTA	>TS	V GTA
20	R AGA	R AGA	R: AGA	R AGA		A GCA	A GCA	GCA	
	M ATG	M ATG	ATA	${\sf ATA}$		6 66A	6 66A	V GTA GGA	6 A 66A 6CA
	R AGA	R AGA	R AGA	R AGA		۷ 6T6	۷ 6T6	V GTA	V GTA
	٠								
	LAV.BRU	ARV 2	LAV.MAL	LAV.ELI		LAV.BRU	ARV 2	LAV.MAL	LAV, EL I
ပ	_	A		_	Q	آ	A	Ĺ	Ī

FIG. 6A-3

EN

ð

CTC LAV'BRU CAG CAC TTG TGG ACA TGG GGC TGG AAA TGG GGC ACC ATG

20

L CTC L TT6 ACC TE TEG AGA TEG GEC CAC Q CAG ARV 2

L CTC M ATG M ATG LAV.MAL CAA AAC TGG TGG AGA TGG GGC

L CTC LAV.ELI CAA AAC TGG TGG AAA TCG GGC

6B-1

E G

150

- GGG AAT GCT ACT AAT ACC AAT AGT AGT ACC AAT AGT AGT AGT GG GAA GCA GAG ATA TGC ACT GAT TTG

M E K K AAG

140

LAV.BRU

GCA GAG K AAA M ATG

ARV 2

THA AAT TGC ACT GAT TTG - GGG AAG GCT ACT AAT ACC AAT AGT AGT

W K E E E I K G E I

TGG AAA GAA GAA ATA AAA GGA GAA ATA

LAV.MAL

THA AAC TGC ACT AAT GTG AAT GGG ACT GTG AAT GGG ACT AAT GCT GGG AGT AAT AGG ACT AAT GCA GAA

L K M E I G E V TTG AAA ATG GAA ATT - GGA GAA GTG

LAV.ELI

THE AGG AAC AAT GGC ACT ATG GGG AAC AAT GTC ACT ACA GAG GAG AAA TTA AAC TGT AGT GAT GAA

FIG. 6B-2

Y T L TAT ACG TTG Y R L TAT AGG CTA GAT AAT GCT AGT ACT ACT ACC AAC TAT ACC AAC TAT AGG TTG B N D S S T N S T N Y R L SAC AAT GAT AGT ACC AAT TAT AGG TTA GAT GAT AGT GAT AAT AGT AGT - - -D N D T T S GAT AAT GAT ACT ACC AGC -LAV, BRU LAV, ELI LAV, MAL ARV 2

6B-3

LAV, BRU

TGT AAT TCA ACA CAA CTG TTT AAT AGT TTT AAT AGT ACT TGG AGT ACT GAA GGG TCA AAT AAC ACT GAA GGA AGT ACT AGT GAC ACA AT AAC ACT GAA GGA AGT GAC ACA ATC

ARV 2

AGG TTA AAT CAC ACT GAA GGA ACT AAA GGA C N T T Q L F N N T W T TGT TGT AAT ACA TGG

D T I GAC ACA ATC

LAV.MAL

S N S T E SA AGT AGT AGT TCA

LAV, EL I

TGT AAT ACA TCA GGA CTG TTT AAT AGT ACA TGG AAT ATT AGT GCA TGG AAT AAT ATT ACA GAG TCA AAT AAT AGC ACA N T N I AAC ACA AAC ATC

FIG. 6B-4

LAV. MAL

R |GGTCTCTCTTGTTAGACCAGGTCGAGCCCGGGAGCTCTCTGGCTAGCAAGGAACCCACTG CTTAAGCCTCAATAAAGCTTGCCTTGAGTGCCTCAAGCAGTGTGCCCATCTGTTGTGT GACTCTGGTAACTAGAGATCCCTCAGACCACTCTAGACGGTGTAAAAATCTCTAGCAGTG GCGCCCGAACAGGGACTTTAAAGTGAAAGTAACAGGGACTCGAAAGCGGAAGTTCCAGAG AAGTTCTCTCGACGCAGGACTCGGCTTGCTGAGGTGCACACAGAGAGGCGAGAGCGGC r. →GAG 300. AlaSerValLeuSerGlyĠlyLysLeuAspAlaTrpGluLysIleArgLeuArgProGly AGCGTCAGTATTAAGCGGGGGAAAATTAGATGCATGGGAGAAAATTCGGTTAAGGCCAGG GlyLysLysLysTyrArgLeuLysHisLeuValTrpAlaSerArgGluLeuGluArgPhe GGGAAAGAAAAATATAGACTGAAACATTTAGTATGGGCAAGCAGGGAGCTGGAAAGATT AlaLeuAsnProGlyLeuLeuGluThrGlyGluGlyCysGlnGlnIleMetGluGlnLeuCGCACTTAACCCTGGCCTTTTAGAAACAGGAGAAGGATGTCAACAAATAATGGAACAGCT 500 GlnSerThrLeuLysThrGlySerGluGluIleLysSerLeuTyrAsnThrValAlaThr ACAATCAACTCTCAAGACAGGATCAGAAGAAATTAAATCATTATAATACAGTAGCAAC LeuTyrCysValHisGlnArgIleAspValLysAspThrLysGluAlaLeuAspLysIle CCTCTATTGTGTACATCAAAGGATAGATGTAAAAGACACCAAGGAAGCGCTAGATAAAAT AlaAlaAlaThrLysAsnSerSerSerValSerGlnAsnTyrProIleValGlnAsnAla AGCAGCTGCCACAAAAAACAGCAGCAGTGTCAGTCAAAATTACCCCATAGTGCAAAATGC GlnGlyGlnMetIleHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysVal ACAAGGGCAAATGATACATCAGGCCATATCACCTAGGACTTTGAATGCATGGGTGAAAGT 800 IleGluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGiy AATAGAAGAAAAGGCTTTCAGCCCAGAAGTGATACCCATGTTCTCAGCATTATCAGAGGG AlaThrProGlnAspLeuAsnMetMetLeuAsnIleVaiGlyGlyHisĠlnAlaAlaMet GGCCACCCCACAAGATTTAAATATGATGCTGAACATAGTTGGAGGACACCAGGCAGCTAT GlnMetLeuLysAspThrileAsnGluGluAlaAlaAspTrpAspArgValHisProValGCAAATGTTAAAAGATACCATCAATGAGGAAGCTGCAGACTGGGACAGGGTACATCCAGT 1000 HisAlaGlyProlleProProGlyGlnMetArgGluProArgGlySerAsplleAlaGly ACATGCAGGGCCTATTCCCCCAGGCCAGATGAGAGAACCAAGAGGAAGTGACATAGCAGG

ThrThrSerThrLeuGlnGluGlnIleGlyTrpMetThrSerAsnProProIleProValAACTACTAGTACCCTTCAAGAACAAATAGGATGGATGACAAGCAACCCACCTATCCCAGT 1100 1200 ProvalSerIleLeuAspileArgGlnGiyProLysGluProPheArgAspTyrValAsp CCCTGTCAGCATTTTGGACATAAGACAAGGGCCAAAGGAACCTTTTAGAGACTATGTAGA ArgPhePheLysThrLeuArgAlaGluGİnAlaThrGlnGluValLysAsnTrpMetThr TAGGTTCTTTAAAACTCTCAGAGCTGAGCAAGCTACACAGGAGGTAAAAAATTGGATGAC 1300 GluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGiy AGAAACCTTGCTGGTCCAAAATGCGAATCCAGACTGTAAGACCATTTTAAAAGCATTAGG ProGlyAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProSerHis ACCAGGGGCTACATTAGAAGAAATGATGACAGCATGCCAGGGAGTGGGAGGACCCAGTCA 1400 LysAlaArgValLeuĀlaGluAlaMetSerGlnAlaThrAsnSerThrAlaAlaIleMet TAAAGCAAGAGTTTTGGCTGAGGCAATGAGCCAAGCAACAAATTCAACTGCTGCCATAAT 1500 MetGlnArgGlyAsnPheLysGlyGlnLysArgIleLysCysPheAsnCysGlyLysGluGATGCAGAGAGTAATTTTAAGGGCCAGAAAAGAATTAAGTGTTTCAACTGTGGCAAAGA GlyHisLeuAlaArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLys AGGACACCTAGCCAGAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGGAA ,ŏ r>₽ŎL 1600 GluGlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrp GGAAGGACACCAAATGAAAGACTGCACTGAGAGACAGGCTAATTTTTTAGGGAAAATTTG AlaPheProĠlnGlyLysAlaArgGluPheProSerGluĠlnThrArgAlaAsnSerPro ProSerHisLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaPro GCCTTCCCACAAGGGAAGGCCAGGGAATTTCCTTCAGAGCAGACCAGAGCCAACAGCCCC 1700 ThrSerArgĠluLeuArgValTrpGlyGlyAspLysThrLeuSerGluThrGlyAlaGlu ProAlaGluSerPheGlyPheGlyGluGluIleLysProSerGlnLysGlnGluGlnLys ACCAGCAGAGAGCTTCGGGTTTGGGGAGGAGATAAAACCCTCTCAGAAACAGGAGCAGAA ArgGlnGlyİleValSerPheSerPheProGlnIleThrLeuTrpGlnArgProValVal AspLysGluLeuTyrProLeuAlaSerLeuLysSerLeuPheGlyAsnAspGlnLeuSer AGACAAGGAATTGTATCCTTTAGCTTCCCTCAAATCACTCTTTGGCAACGACCAGTTGTC GAG ← . ThrValArgValGlyGlyGinLeuLysGluAlaLeuLeuAspThrGlyAlaAspAspThr Gln ACAGTAAGAGTAGGAGGACAGCTAAAAGAAGCTCTATTAGACACAGGAGCAGATGATACA ValLeuGluGluIleAsnLeuProGlyLysTrpLysProLysMetIleGlyGlyIleGlyGTATTAGAAGAAATAAATTTGCCAGGAAAATGGAAACCAAAAATGATAGGGGGAATTGGA GlyPheIleLysValArgGinTyrAspGinIleLeuIleGluIleCysGiyLysLysAlaGGTTTTATCAAAGTAAGACAGTATGATCAAATACTTATAGAAATTTGTGGAAAAAAAGGCT 2000

IleGlyThrIleLeuValGlyProThrProValAsnIleIleGlyArgAsnMetLeuThr ATAGGTACAATATTGGTAGGACCTACACCTGTCAACATAATTGGACGAAATATGTTGACT GlnIleGlyĊysThrLeuAsnPheProIleSerProIleGluThrValProValLysLeuCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAAAATTA 2300 LysLeuValAsnPheArgGluLeuAsnLysArgThrGlnAspPheTrpGluValGlnLeu AAATTAGTGAATTTCAGAGAGCTTAATAAAAGAACTCAAGATTTTTGGGAAGTTCAATTA GlyIleProHisProAlaGlyLeuLysLysLysLysSerValThrValLeuAspValGlyGGAATACCACATCCTGCTGGGTTGAAAAAAGAAAAATCAGTCACAGTATTGGATGTGGG 2400 AspAlaTyrPheSerValProLeuAspGluAspPheArgLysTyrThrAlaPheThrIle GATGCATATTTTCAGTCCCTTTAGATGAAGATTTCAGGAAGTATACTGCATTCACTATA ProSerIleAsnAsnGluThrProGlyIleArgTyrGlnTyrAsnValLeuProGlnGlyCCCAGTATTAATAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTACCACAGGGA TrpLysGlySerProAlaIlePheGlnSerSerMetThrLysIleLeuGluProPheArgTGGAAAGGATCACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAACCCTTTAGA 2700 LeuGluIleĠlyGlnHisArgThrLysIleGluGluLeuArgGluHisLeuLeuLysTrp TTAGAAATAGGACAACATAGAACAAAAATAGAGGAACTAAGAGAACATCTATTGAAATGG GlyPheThrThrProAspLysLysHisGlnLysGluProProPheLeuTrpMetGlyTyr GGATTTACCACACCAGACAAAAAGCATCAGAAAGAACCCCCCATTTCTTTGGATGGGGTAT GluLeuHisProAspLysTrpThrValGlnProIleGlnLeuProAspLysGluSerTrp GAACTCCACCCTGACAAATGGACAGTGCAGCCTATACAACTGCCAGACAAGGAAAGCTGG ThrValAsnAspIleGlnLysLeuValGlyLysLeuAsnTrpAlaSerGlnIleTyrPro ACTGTCAATGATATACAGAAATTGGTGGGAAAACTAAATTGGGCAAGTCAGATTTATCCA 2900 GlyIleLysValLysGInLeuCysLysLeuLeuArgGlyAlaLysAlaLeuThrAspIle GGAATTAAAGTAAAGCAATTATGTAAACTCCTTAGGGGAGCAAAAGCACTAACAGACATA ValProLeuThrAlaGluAlaGluLeuGluLeuAlaGluAsnArgGluIleLeuLysGlu GTACCATTAACTGCAGAGGCAGAATTAGAATTGGCAGAGAACAGGGAAATTCTAAAAGAA

ProvalHisGlyValTyrTyrAspProSerLysAspLeuIleAlaGluIleGlnLysGlnCCAGTGCATGGGGTATATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGGGGGCAAGGTCAATGGACATATCAAATATACCAAGAGCAATATAAAAAATCTGAAAAACAGGG LysTyrAlaArgIleLysSerAlaHisThrAsnAspValLysGlnLeuThrGluAlaVal AAGTATGCAAGAATAAAGTCTGCCCACACTAATGATGTAAAACAATTAACAGAAGCAGTG GlnLysIleAlaGlnGluSerIleValIleTrpGlyLysThrProLysPheArgLeuProCAAAAGATAGCCCAAGAAAGCATAGTAATATGGGGAAAAACTCCTAAATTTAGACTACCC 3300 IleGlnLysĠluThrTrpGluAlaTrpTrpThrGluTyrŤrpGlnAlaThrTrpIlePro ATACAAAAAGAAACATGGGAGGCATGGTGGACAGAATATTGGCAAGCCACCTGGATCCCT GluTrpGluPheValAsnThrProProLeuValLysLeuTrpTyrGlnLeuGluThrGluGAATGGGAGTTTGTCAATACTCCTCCCCTAGTAAAACTATGGTACCAGTTAGAAACAGAA 3400 ProllevalglyAlaGluThrPheTyrValAspGlyAlaAlaAsnArgGluThrLysLysCCCATAGTAGGAGCAGAAACTTTCTATGTAGATGGGGCAGCTAATAGAGAAACTAAAAAG GlyLysAlaĠlyTyrValThrAspArgGlyArgGlnLysValValSerLeuThrGluThr GGAAAAGCAGGATATGTTACTGACAGAGGAAGACAAAAGGTTGTCTCCTTAACTGAAACA 3500 ThrAsnGlnLysThrGluLeuGlnAlaIleHisLeuAlaLeuGlnAspSerGlySerGluACAAATCAGAAGACTGAATTACAAGCAATCCACTTAGCTTTACAGGATCAGAA 3600 ValAsnIleValThrAspSerGlnTyrAlaLeuGlyIleİleGlnAlaGlnProAspLys GTAAACATAGTAACAGACTCACAGTATGCATTAGGGATTATTCAAGCACAACCAGATAAA SerGluSerĠluIleValAsnGlnIleIleGluGlnLeuİleGlnLysAspLysValTyr AGTGAATCAGAGATTGTTAATCAAATAATAGAGCAATTAATACAGAAGGACAAGGTCTAC LeuSerTrpValProAlaHisLysGlyIleGlyGlyAsnGluGlnValAspLysLeuValCTGTCATGGGTACCAGCACACAAGGGATTGGAGGAAATGAACAAGTAGATAAATTAGTC GlnValAspCysSerProGlyIleTrpGlnLeuAspCysThrHisLeuGluGlyLysIleCAAGTAGACTGTAGTCCAGGGATATGGCAATTAGATTGCACACATCTAGAAGGAAAAATA IleIleValAlaValHisValAlaSerGlyTyrIleGluAlaGluValIleProAlaGlu ATCATAGTAGCAGTCCATGTAGCCAGTGGATATATAGAAGCAGAAGTTATCCCAGCAGAA ThrGlyGlnGluThrAlaTyrPheIleLeuLysLeuAlaGlyArgTrpProValLysValACAGGACAGGAGACAGCATACTTTATACTAAAATTAGCAGGAAGATGGCCAGTAAAAGTA 4100

ValHisThrAspAsnGlySerAsnPheThrSerAlaAlaValLysAlaAlaCysTrpTrpGTACACACAGACAATGGCAGCAATTTCACCAGTGCTGCAGTTAAAGCAGCCTGTTGGTGG 4200 AlaAsnIleLysGlnGluPheGlyIleProTyrAsnProGlnSerGlnGlyValValGluGCAAATATCAAACAGGAATTTGGAATTCCCTACAACCCCCAAAGTCAAGGAGTAGTGGAA SerMetAsnLysGluLeuLysLysIleIleGlyGlnValArgGluGlnAlaGluHisLeu TCTATGAATAAGGAATTAAAGAAAATCATAGGGCAGGTAAGAGAGCAAGCTGAACACCTT 4300 LysThrAlavalGlnMetAlavalPheIleHisAsnPheLysArgLysGlyGlyIleGlyAAGACAGCAGTACAAATGGCAGTGTTCATTCACAATTTTAAAAAGAAAAGGGGGGATTGGG GlyTyrSerAlaGlyGluArgIleIleAspMetIleAlaThrAspIleGinThrLysGluGGGTACAGTGCAGGGGAAAGAATAATAGACATGATAGCAACAGACATACAAACTAAAGAA 4400 LeuGlnLysGlnIleThrLysIleGlnAsnPheArgValTyrTyrArgAspAsnArgAsp TTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAACAGAGAC 4500 ProlleTrpLysGlyProAlaLysLeuLeuTrpLysGlyGluGlyAlaValValIleGlnCCAATTTGGAAAGGACCAGCAAAACTACTCTGGAAAGGTGAAGGGGCAGTAGTAATACAG AspAsnSerAspIleLysValValProArgArgLysAlaLysIleIleArgAspTyrGly MetGlu GACAATAGTGATATAAAGGTAGTACCAAGAAGAAAAGCAAAAATCATTAGGGATTATGGA 4600 POL← LysGlnMetAlaGlyAspAspCysValAlaGlyGlyGlnAspGluAsp
AsnArgTrpGlnValMetIleValTrpGlnValAspArgMetArgIleArgThrTrpHis
AAACAGATGGCAGGTGATGATTGTGTGGCAGGTGGACAGGATGAGGATTAGAACATGGCA SerLeuValLysHisHisMetTyrValSerLysLysAlaLysAsnTrpPheTyrArgHisCAGTTTAGTAAAACATCATATGTATGTCTCAAAGAAAGCTAAAAATTGGTTTTATAGACA 4700 HISTYrGluSerArgHÍSProLysValSerSerGluValHISIleProLeuGlyAspAla TCACTATGAAAGCAGGCATCCAAAAGTAAGTTCAGAAGTACACATCCCACTAGGGGATGC 4800 ArgLeuValValArgThrTyrTrpGlyLeuGlnThrGlyGluLysAspTrpHisLeuGly TAGATTAGTAGTAAGAACATATTGGGGTCTGCAAACAGGAGAAAAAGACTGGCACTTGGG HisGlyValSerIleGluTrpArgGlnLysArgTyrSerThrGlnLeuAspProAspLeuTCATGGGGTCTCCATAGAATGGAGGCAGAAAAGATATAGCACACAACTAGATCCTGACCT AlaAspGlnLeuIleHisLeuTyrTyrPheAspCysPheSerGluSerAlaIleArgGin AGCAGACCAACTGATTCATCTGTACTATTTTGATTGTTTTTCAGAATCTGCCATAAGACA AlaIleLeuGlyHisIleValSerProArgCysAspTyrGlnAlaGlyHisAsnLysVal AGCCATATTAGGACATATAGTTAGTCCTAGGTGTGATTATCAAGCAGGACATAACAAGGT 5000 GlySerLeuGlnTyrLeuAlaLeuThrAlaLeuIleAlaProLysLysThrArgProProAGGATCTTTACAGTATTTGGCACTAACAGCATTAATAGCACCAAAAAAGACAAGGCCACC **→**R MetGluGlnAlaProAlaAspGlnGlv LeuProSerValArgLysLeuThrGluAspAngTrpAsnLysProGlnGlnThrLysGly TTTGCCTAGTGTTAGGAAGCTAACAGAAGATAGATGGAACAAGCCCCAGCAGACCAAGGG

ProglnArgGluProHisAsnGluTrpThrLeuGluLeuLeuGluGluLeuLysGlnGluHisArgGlySerHisThrMetAsnGlyHisCCACAGAGGGAGCCACACAATGAATGGACATTAGAACTTTTAGAGGAGCTTAAGCAAGAA 5200 AlaValArgHisPheProArgIleTrpLeuHisSerLeuGlyGlnHisIleTyrGluThr GCTGTCAGACACTTTCCTAGGATATGGCTCCATAGTTTAGGACAACATATCTATGAAACT TyrGlyAspThrTrpGluGlyValGluAlaIleIleArgSerLeuGlnGlnLeuLeuPhe TATGGGGATACCTGGGAAGGAGTTGAAGCTATAATAAGAAGTCTGCAACAACTGCTGTTT 5300 IleHisPheArgIleGlyCysGlnHisSerArgIleGlyİleThrArgGlnArgArgAla ATTCATTTCAGAATTGGGTGTCAACATAGCAGAATAGGCATTACTCGACAGAGAAGAGCA R← 5400 ArgAsnGlySerSerArgSer MetAspProValAspProAsnLeuGluProTrpAsnHisProGlySerGlnProArg AGAAATGGATCCAGTAGATCCTAACTTAGAGCCCTGGAACCATCCAGGGAGTCAGCCTAG ThrProCysAsnLysCysTyrCysLysLysCysCysTyrHisCysGlnMetCysPhelie GACGCCTTGTAATAAGTGTTATTGTAAAAAGTGCTGCTATCATTGCCAAATGTGCTTCAT 5500 ThrLysGlyLeuGlyIleSerTyrGlyArgLysLysArgArgGlnArgArgArgProProAACGAAAGGCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGACCTCC GTGGACCATÁGTATTTATAĠAAATTAGGAÁAATAAGAAGÁCAAAGGAAAÁTAGACAGGTT **→**ENV GATTGATAGAATAAGAGAAGAGAGAGAGATAGTGGCAATGAGAGTGAGGGAGATACAGA 5800 AsnTyrGlnAsnTrpTrpArgTrpGlyMetMetLeuLeuGlyMetLeuMetThrCysSer GGAATTATCAAAACTGGTGGAGATGGGGCATGATGCTCCTTGGGATGTTGATGACCTGTA IleAlaGluAspLeuTrpValThrValTyrTyrGlyValProValTrpLysGluAlaThr GTATTGCAGAAGATTTGTGGGGTTACAGTTTATTATGGGGTACCTGTGTGGAAAGAAGCAA 5900 ThrThrLeuPheCysAlaSerAspAlaLysSerTyrGluThrGluValHisAsnIleTrp CCACTACTCTATTTTGTGCATCAGATGCTAAATCATATGAAACAGAAGTACATAACATCT AlaThrHisAlaCysVaiProThrAspProAsnProGinGluIleGluLeuGluAsnValGGGCTACACATGCCTGTGTACCCACGGACCCCACACCCACAGAAATAGAACTGGAAAATG ThrGluGlyPheAsnMetTrpLysAsnAsnMetValGluGlnMetHisGluAspIleile TCACAGAAGGGTTTAACATGTGGAAAAATAACATGGTGGAGCAGATGCATGAGGATATAA

SerLeuTrpAspGlnSerLeuLysProCysValLysLeuThrProLeuCysValThrLeuTcAGTTTATGGGATCAAAGCCTAAAACCATGTGTAAAGCTAACCCCACTCTGTGTCACTT AsnCysThrAsnValAsnGlyThrAlaValAsnGlyThrAsnAlaGlySerAsnArgThr TAAACTGCACTAATGTGAATGGGACTGCTGTGAATGGGACTAATGCTGGGAGTAATAGGA 6200 AsnAlaGiuLeuLysMetGluIleGlyĠluValLysAsnCysSerPheAsnIleThrÞro CTAATGCAGAATTGAAAATGGAAATTGGAGAAGTGAAAAACTGCTCTTTCAATATAACCC ValGlySerAspLysArgGlnGluTyrAlaThrPheTyrAsnLeuAspLeuValGlnIle CAGTAGGAAGTGATAAAAGGCAAGAATATGCAACTTTTTATAACCTTGATCTAGTACAAA AspAspSerAspAsnSerSerTyrArgLeuIleAsnCysAsnThrSerValIleThrGln TAGATGATAGTGATAATAGTAGTTATAGGCTAATAAATTGTAATACCTCAGTAATTACAC 6400 AlaCysProLysValThrPheAspProileProileHisTyrCysAlaProAlaGlyPheAGGCTTGTCCAAAGGTAACCTTTGATCCAATTCCCATACATTATTGTGCCCCAGCTGGTT AlaIleLeuLysCysAsnAspLysLysPheAsnGlyThrGluIleCysLysAsnValSerTTGCAATTCTAAAGTGTAATGATAAGAAGTTCAATGGAACGGAAATATGTAAAAATGTCA ThrValGinCysThrHisGlyIleLysProValValSerThrGlnLeuLeuLeuAsnGlyGTACAGTACAATGTACACATGGAATTAAGCCAGTGGTGTCAACTCAACTGCTGTTAAATG SerLeuAlaGluGluIleMetIleArgSerGluAsnLeuThrAspAsnThrLysAsn GCAGTCTAGCAGAAGAAGAGATAATGATTAGATCTGAAAATCTCACAGACAATACTAAAA IleIleValGlnLeuAsnGluThrValThrIleAsnCysThrArgProGlyAsnAsnThrACATAATAGTACAGCTTAATGAAACTGTAACAATTAATTGTACAAGGCCTGGAAACAATA 6700 ArgArgGİyIleHisPheGlyProGlyGlnAlaLeuTyrThrThrGlyIleValGlyAsp CAAGAAGAGGGATACATTTCGGCCCAGGGCAAGCACTCTATACAACAGGGATAGTAGGAG IleArgArgAlaTyrCysThrIleAsnGluThrGluTrpAspLysThrLeuGlnGlnValATATAAGAAGAGCATATTGTACTATTAATGAAACAGAATGGGATAAAACTTTACAACAGG AlaValLysLeuGlySerLeuLeuAsnLysThrLysIleIlePheAsnSerSerSerGly TAGCTGTAAAACTAGGAAGCCTTCTTAACAAAACAAAAATAATTTTTAATTCATCCTCAG 6900 GlyAspProGluIleThrThrHisSerPheAsnCysArgGlyGluPhePheTyrCysAsnGAGGGGACCCAGAAATTACAACACACAGTTTTAATTGTAGAGGGGAATTTTTCTACTGTA ThrSerLysLeuPheAsnSerThrTrpGlnAsnAsnGlyAlaArgLeuSerAsnSerThrAraCATCAAAACTGTTTAATAGTACATGGCAGAATAATGGTGCAAAACTAATAGTACATAGCA GluSerThrGlySerIleThrLeuProCysArgIleLysGlnIleIleAsnMetTrpGlnCAGAGTCAACTGGTAGTATCACACTCCCATGCAGAATAAAACAAATTATAAATATGTGGC LysThrGlyLysAlaMetTyrAlaProProIleAlaGlyValIleAsnCysLeuSerAsnAGAAAACAGGAAAAGCTATGTATGCCCCTCCCATCGCAGGAGTCATCAACTGTTTATCAA 7100

I leThrGlyLeuI leLeuThrArgAspGlyGlyAsnSerSerAspAsnSerAspAsnGlu ATATTACAGGGCTGATATTAACAAGAGATGGTGGAAATAGTGACAATAGTGACAATG

7200

ThrLeuArgProGlyGlyGlyAspMetArgAspAsnTrpIleSerGluLeuTyrLysTyrAGACCTTAAGACCTGGAGGAGGAGATATGAGGGACAATTGGATAAGTGAATTATAAAT GluArgGluLysArgAlaIleGlyLeuGlyAlaMetPheLeuGlyPheLeuGlyAlaAla TGGAAAGAGAAAAAAGAGCAATAGGACTAGGAGCCATGTTCCTTGGGTTCTTGGGAGCAG GlySerThrMetGlyAlaAlaSerLeuThrLeuThrValGlnAlaArgGlnLeuLeuSerCAGGAAGCACGATGGGCGCAGCGTCACTAACGCTGACGGTACAGGCCAGACAGTTACTGT 7400 GlyIleValGlnGlnGlnAsnAsnLeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeu CTGGTATAGTGCAACAGCAAAACAATTTGCTGAGGGCTATAGAGGCGCAACAGCATCTGT GlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgValLeuAlaValGluÁrgŤyr TGCAACTCACGGTCTGGGGCATTAAACAGCTCCAGGCAAGAGTCCTGGCTGTGGAAAGAT LeuGlnAspGlnArgLeuLeuGlyMetTrpGlyCysSerGlyLysHisIleCysThrThr ACCTACAGGATCAACGGCTCCTAGGAATGTGGGGTTGCTCTGGAAAACACATTTGCACCA 7600 PheValProTrpAsnSerSerTrpSerAsnArgSerLeuAspAspIleTrpAsnAsnMet CATTTGTGCCTTGGAACTCTAGTTGGAGTAATAGATCTCTAGATGACATTTGGAATAATA ThrTrpMetGlnTrpGluLysGluIleSerAsnTyrThrGlyIleIleTyrAsnLeuile TGACCTGGATGCAGTGGGAAAAAGAAATTAGCAATTACACAGGCATAATATACAACTTAA 7700 GluGluSerGlnIleGlnGlnGluLysAsnGluLysGluLeuLeuGluLeuAspLysTrp TTGAAGAATCGCAAATCCAGCAAGAAAAGAATGAAAAGGAATTATTGGAATTGGACAAGT AlaSerLeuTrpAsnTrpPheSerIleSerLysTrpLeuTrpTyrIleArgIlePheIle GGGCAAGTTTGTGGAATTGGTTTAGCATATCAAAATGGCTGTGGTATATAAGAATATTCA ArgValArgGlnGlyTyrSerProLeuSerLeuGlnThrLeuLeuProThrProArgGly ATAGAGTTAGGCAGGGATACTCACCTCTGTCGTTGCAGACCCTCCTCCCAACACCGAGGG ProProAspArgProGluGlyIleGluGluGluGlyGlyGluGlnGlyArgGlyArgSer GACCACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGCAAGGCAGAGGCAGAT 8000 IleArgLeuValAsnGlyPheSerAlaLeuIleTrpAspAspLeuArgAsnLeuCysLeuCAATTCGATTGGTGAACGGATTCTCAGCACTTATCTGGGACGACCTGAGGAACCTGTGCC PheSerTyrHisArgLeuArgAspLeuLeuLeuIleAlaThrArgIleValGluLeuLeu TCTTCAGTTACCACCGCTTGAGAGACTTACTCTTAATTGCAACGAGGATTGTGGAACTTC GlyArgArgGlyTrpGluAlaLeuLysTyrLeuTrpAsnLeuLeuGlnTyrTrpGlyGlnTGGGACGCAGGGGGGGAAGCCCTCAAATATCTGTGGAATCTCCTGCAATATTGGGGTC 8200

FIG. 7H

GluLeuLysAsnSerAlaIleSerLeuLeuAsnThrThrAlaIleAlaValAlaGluCys AGGAACTGAAGAATAGTGCTATTAGCTTGCTTAATACCACAGCAATAGCAGTAGCTGAAT ThrAspArgValIleGluIleGlyGlnArgPheGlyArgAlaIleLeuHisIleProArg GCACAGATAGGGTTATAGAAATAGGACAAAGATTTGGTAGAGCTATTCTCCACATACCTA 8300 MetGlyGlyLysTrpSerLys EW~ ArgIleArgGlnGlyPheGluArgAlaLeuLeu GAAGAATTAGACAGGGCTTCGAAAGGGCTTTGCTATAACATGGGTGGCAAGTGGTCAAAA 8400 SerSerIleValGlyTrpProLysIleArgGluArgIleArgArgThrProProThrGlu AGTAGCATAGTAGGATGGCCTAAGATTAGGGAAAGAATAAGACGAACTCCCCCAACAGAA ThrGlyValGlyAlaValSerGlnAspAlaValSerGlnAspLeuAspLysCysGlyAlaACAGGAGTAGGAGCAGTATCTCAAGATGCAGTATCTCAAGATTTAGATAAATGTGGAGCA GluValGlyPheProValArgProGlnValProLeuArgProMetThrTyrLysGlyAla GAGGTAGGCTTTCCAGTCCGTCCTCAGGTACCTTTAAGACCAATGACTTATAAAGGAGCT PheAspLeuSerHisPheLeuLysGluLysGlyGlyLeuAspGlyLeuValTrpSerPro TTTGATCTCAGCCACTTTTTAAAAGAAAAGGGGGGGACTGGATGGGTTAGTTTGGTCCCCA GlnAsnTyrThrProGlyProGlyIleArgPheProLeuThrPheGlyTrpCysPheLysCAGAATTACACACCAGGGCCAGGGATTAGATTCCCACTGACCTTCGGATGGTGCTTTAAG LeuValProMetSerProGluGluValGluGluAlaAsnGluGlyGluAsnAsnCysLeuTTAGTACCAATGAGTCCAGAGGAAGTAGAGGAGGCCAATGAAGGAGAACAACTGTCTG 8900 PheAspSerSerLeuAlaLeuArgHisArgAlaArgGluĠlnHisProGluTyrTyrLys TTTGACAGCAGCCTAGCACTAAGACACAGAGCCAGAGAACATCCGGAGTACTACAAA F← 9000 ASDCYS GACTGCTGACAGAGGTTGCTGACAGGGGACTTTCCGCTGGGGACTTTCCAGGGGAGGC GTAACTTGGGCGGGACCGGGGAGTGGCTAACCCTCAGATGCTGCATATAAGCAGCTGCTT TTCGCCTGTACTGGGTCTCTCTTGTTAGACCAGGTCGAGCCCGGGAGCTCTCTGGCTAGC U3<-->R AAGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCCTCAA

FIG. 1B FIG. 1A LAV eli LAV mal LAV bru **↓"*** LAV mal ٥٥ Į,

FIG. 2 LAV mal LAV bru

				•
y. 4 0	CSELLNSL T QAISPRTL	200 210 220 230 240 RTHLNTVGCH QAAHQNLKET INEEAAEUDK VHFVHACPIA FGUREPRGS H I D D L	300 310 320 149CPKEPFR DYVDRFYKTL RAEGASGEVK F T D 500 \$\\$750 \	RPEPTAPPE 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
7.0	LETSECCRQ1 LCQLQFSLQT CSELLKSLYN C Q HE ST K IK D 25	230 VHF VHACF LA F	310 by vdrftktl F F 1990 ATTHHQRGHF H A T A	EPTAPIFIQS
g	LETSECCRQ1 LCQ1 C Q HE K 1 D 25 150 S S Q V S Q HY F I V Q HI S S	220 1reeaaeudk D		APCHFLQSAP
Ş	LERFAVIPCL Y L 130 AC R AQQAAA KH	210 Qaahqiilket D	290 RHYSPTSILD V V 370 GPCHKARVLA S	LCKIVPSYKG H H
07	LKIII VVA SKE L 120 SKKKAQQAAA AQ T	200 RIMLNIVGGH M 1	280 W111CLHKIV V 360 EMHTACQGVG	440 KDCTERQAHF LFGNDYSSQ QL L
	CELDRUEKIR LRPCCKKYYK LKHIVVASKE LEKFAVHPCL LETSECCRQI LCQLQFSLQT K A	190 LSECATPQDL		WKCCKECHQH KDCTERQAHF R R L S 10 S 10 ELYPLTSLKS LFCHDPSSQ A K QL L
	CELDRUEKIN K K K K 100 RIEIKDTKEA DV C DV	180 Speviphesa	200 270 EQICUITIIIP PIPVCEIYKR S S D S D S D S D D S D D S D D D S D D D S D	420 ncrapresco 500 psqrqepide qr qr
	HCARMSVLSC 90 TVATLYCVHQ 1	UVKVVEEKAF	150 DIACTTSTLQ 330 RUNT ETLL VQ	410 NCGEEGHAR L L L L A 490 SFRSCVETIT F E K CF E I
G A G	LAV BRU LAV MAL LAV BRU ARV 2 LAV BRU LAV BRU LAV BRU	LAV BRU ARV 2 LAV MAL LAV EL1	LAV BRU ARY 2 LAV 4 LAV HAL LAV BRU ARV 3 LAV HAL LAV HAL	LAV BRU ARV 2 LAV HAL LAV EL1 LAV EL1 ARV 2 LAV HAL LAV EL1

DSCIISSE CELESE

80 ITLUQRPLVI V	160 H11CRELLTQ H H KKKDSTKWAK	NGVIIVLEGGU 400 400	AENREILKEP SOCKTPKFKLPI I R
70 QCTVSFNFPQ 1 S	140 150 LIEICGHAAI GTVLVGPTPV P	310 Simetecia 390 Litedakuqk F	A DIV A DVKQLTEAVQ KITTESIVIM CKTPKFKLPI A R S R R A R S R R A R S R R A R R
60 MSLSEAGADA KT T E P KT E	140 LIEICGHKA1 PV K P Q 220 EGKISKIGFE	FAKTAFIIP S S S S S S S S S S S S S S S S S S	A DIV A DIV A DIV A A
AELQVUGRDH MSLSEAGADR QCIVSFIIFPQ ITLUGRFLVI Ge R G - KT T E 1 S V R - P KT E A	FIKVRQYDQI LIEICCHKA1 FY K R P Q 220 LVEICTEHEK ECKISKIGPE T KD L	AYFSVPLDED AYFSVPLDED	LINMASQIYPG IKWRQLCKLL A K K K K K K A K K K K K K K K K K
EQT # 40 40 40 40 40 40 40 40 40 40 40 40 40	METEEKIKA LVEICTEHEK T KD T KD	270 280 IPHPACLKKK KSVTVLDVCD QHPDIVIYQY HDDLYVCSDL K E EH	430 440 450 WIDIQKLVGK LNVASQIYPG IKVRQLCKLL A K K N ER \$10 \$20 \$30 QQUIYQIYQ EPFKHLKTGK YARTRGAHTH H QY H H N H N
	110 LEENSLPCRH N K LM K LM K LH K LH K LH K LH K LH K LH K LH K LH	270 IPHPAGLKKK 350 QHPDIVIYQY K E EH	430 WHDIQKLVCK N ER 510 QCQVTYQIYQ
20 GKAREFSSEQ P C L PK	A LEDICADDIV LEENSLPCRW N K IN K N K O 180 180 190 S PIETVPVKLK PCHDGPKVKQ	TQDFWEYQLG IPHPAGLKKK KSVTVLDVCD AYFSYPLDED FAKYTAFTIF SIHHETFGIR YQYHVLPQGU TQDFWEYQLG IPHPAGLKKK KSVTVLDVCD AYFSYPLDED FAKYTAFTIF SIHHETFGIR YQYHVLPQGU S HTKILEFFRK QHYDIVIYQY HDDLYVGSDL EIGQHRYKIE ELRQHLLRWG LTTPDKKHQK EFFFLUHCYE T K E T K E F F F	420 IVLPEKDSUT II Q D E K E S S O D DLIAEIQKQC
FFREDLAFLQ GRAREFSSEQ TRANSPIFSS IN P C L PK	90 IKIGCQLKEA R VRV 170 ICCTLNFPIS	250 LVDFRELHKR H H 330 KGSPAIFQSS	LHPDKWTWQP LHPDKWTWQP S VHCVTYDPSK E
	3 RU 2 2 3 1 A L E L 1 E L 1 2 2 2 2 3 1 A L	8 8 U E L I	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	LARA LARA	

COCLIONS CHICE

10 580 590 600 610 620 630 640 AT EYUQATWIPE WEFVHTPPLV KLWYQIEKEP IVGAETFYVD GAASRETKLG KACYVTKRGR QKVVILIDIT N D S1A A N K D S E	720 GREQVDKLVS	600 DCTHLEGKV1 1	360 KP4SQCVVES	96C RVYYRDSKDP KK BK	
630 KACYVTKRCR D D	710 AUVPAHKG1G S	790 VDCSPGIVQL	B70 CIKQEFCIPY N	950 9 КQІТКІ Q ИГ	
620 GAASRETKLG N N K	650 660 670 680 690 700 710 720 NQKTELQAIN LALQDSGLEV NIVTDSQYAL GIIQAQPDKS ESELVNQIIE QIIKKEKVYL AWVPANKCIG GHEQVDKLVS S S I Q D S	740 750 760 770 780 790 790 790 790 790 790 790 790 80 790 790 80 80 80 80 80 80 80 80 80 80 80 80 80	BIO BIO BIO BIO G40 BSO BAO Abo Lvaviivascy ieaevipaet gqetayfilk lagripvkti htdigshfts tivkaacuna girqefcipy kpqsqcvves I vv aa n Vy aa	900 910 920 930 950 960 950 960 960 960 960 960 960 960 960 960 96	RQDED
610 IVCAETFYVD I	690 ESELVHQIIE S	770 KEIVASCDKC	850 htdhcshfts	930 YSAGERIVDI I H	1010 QHACDDC YAS
600 Kluyqı.ekep T	680 GIIQAQPDKS	760 SDFILLPPVA	640 LAGRUPVKTI VV VY	920 NFKAKGG1GG	1000 KAKIIRDYGK V
590 VEFVNTPPLV	670 HIVTDSQYAL	7 50 KYHSHURAHA H	830 CQETAYFLLK I	910 Tavqhaveiii	980 990 1000 1010 KCECAVVIQD HSDIKVVPRR KAKIIRDYGK QHACDDCVAS RQDED
SBO EYUQATUIPE	660 LALQDSGLEV S	740 CIDKAQDENE E E	820 IEAEVIPAET	900 Qvrdqaeiilk E	980 KCECAVVIQD
\$10 QKETUETUVT \ A H A A	6 50 NQK T E LQA 1 11 K	730 ACIRKVLFLD S 9	8 10 LVAVHVASGY I	M ELEK 11C	970 LWKGPAKLLW 1
BRU 2 HAL EL1	88U 2 11AL EL1	BRU 2 HAL EL1	8 A U 2 4 A L 5 L I	BRU 1 HAL EL1	BAU 2 Mal Eli
1	12 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1	1	12	1 A V L A V

Dedlays ... cales

•	•					
VUATHACVPT I	SSSCENHUE	210 2 220 230 240 THYTH R 1H R T T D T D T D T D T D D T D D T D D T D D T D D T D D T D D T D D T D D T D D T D D T D D T D D D T D D D T D D D T D D D T D D D T D D D T D D D T D D D T D	210 CTRPHHHITRK C C E	390 SECOPLIVI HSFUCCEFF THE TAIL THE THE TAIL THE TAIL THE TAIL THE TAIL THE TAIL THE TAIL THE TAIL TH	480 470 480 480 858117612 LIMDGGHH 5 1 - 4 155	560 CARSHTLTVQ V L A L
JO AKAYDTEVIIN R S C A	ATHTHSSHTH V CTHACS	230 PKVSFEPIPI T D	YQLHQSVEIN E A ET T AH E K T	390 SSCCUPLIVI H T	470 hcsshitcll s h l	S40 TMP SS0 NIFL GFLGAAGSTH H H H
OMP ATTILFGASD	140 SLKCTDL-CN T H N - K T H N Y T	22g CHTSVITQAC R	JOOURTHIE N H	370 380 1ASKLREQFG NNXT-11FKQ V CSLL- K NS V R CSLL- I K P	460 HYAPPISCQ I C A V I ERM	AVGI-CALFL V H V H 1 L- H 1 L- H
50 VYGVP VHKE	130 PCVKLTPLCV	210 TNYTH R 1H TNYTH R 1H S R 1N	290 AEEEVVIRSA IH E	320 IASKLREGYG N VK V CSLL V R GTLL	450 IHHUQEVCKA KT K VAGR-	5307 KKRVVQREKR E
SP 10 10 10 10 10 10 10 10 10 10 10 10 10	110 120 120 130 N	200 111PIDRDITS VV AS T LVQ DSDN V SST	STQLLLNGSL AEEEVVIRSA HFTDHAKTII D W IN E L T M I E L N W	360 RAKWHATLKQ Q H E ETE DK Q	ITLPCRINGF INHUQEVCKA HYAPPISCQI 1	520 530 S40 TMP 550 560 EPLCVAPTKA KRRVVQREKR AVGI-GALFL GFLGAAGSTH GARSHTLTVQ I V H V L V L R E 1 L- H A L V A V
SP 30 CTHLLCILMI	110 HONVEQHHED H Q	190 KEYAFFYKLD M L RI - T H QV L K		MIRGAICHIS RAKWHATLKQ DI K Q H E DI H Y T N ETE DK Q	430 CSHNTEGSDT HTECK W - STGS TES NSTNTN	
20 2MLWRNGUKU NU	100 IVTEHFHHWK G	1STSIRCKVQ T D I TPVCSD R	TCPCTHVSTV QCTHCIREVV K	340 AFVTICK-IC H T RI LY T I-V SLY TKS-RS	420 TUFKSTUSTE RLM Q NCARL- NI A HHI	SOO SIO SIO CCDHRDHURS ELYKYKYVKI I I R R Q Q
HRVKEKY K GTRRK RELQRII	90 DPHPQEVVLV C C 1E E	xCE1KHCSFH - V H	150 LKCHNKTFNC D K	SIKIQACPCA Y Q RIF L Q	YCHSTQLFRS TT M TSK TSC	490 NIGSELFAFC T DT V SDN TL STU T
EN C HAY BRU LAY HAY LLAY ELI	LAV BRU ARY 2 LAV HAL LAV ELI	LAY BRU ARY 2 LAY HAL LAV EL1	LAY BRU ARY 2 LAY HAL LAY EL1	LAV BRU ARV 2 LAV MAL · LAV ELI	LAV BRU ARV 2 LAV HAL LAV ELI	LAY BRU ARY 2 LAY MAL LAY ELI

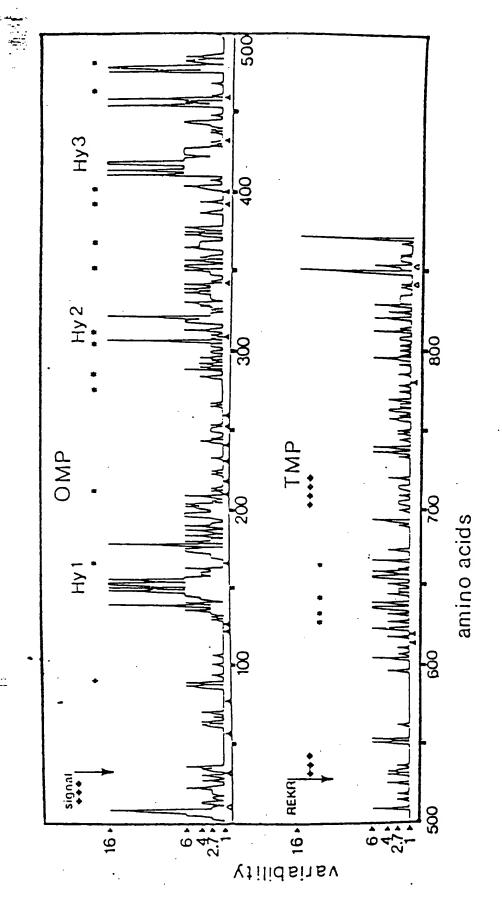
BETED "SZETHOPO

LAV ALL LAV BRU ARV 2 LAV HAL	VLSIVHAVAQ G L L L L BIO RIVELLGRRG I	VDREINNYTS IN EN S G G G G G G G G G G G G G G G G G G	670 XY L YM Y 7 SC PTPRCP-DR V P A - A - C R		LELDKWASLU M PORDRSIALV M QC C V L GC V L BSO 1AVAECTDAV T ECT	700 S S S S S S S S S S S S S S S S S S S	N	CLVCLRIVFA 1
LAV BRU LAV BRU LAV BRU LAV BRU LAV BRU LAV BRU LAV BRU LAV BRU LAV BRU LAV BRU LAV BRU	MCCKVSKSS MCCKVSKSS TPQVPLAPI R R R R R R R E E E E E E E E	10 20 SV VCUPTVRERN H G SA1 1 1 R1 1 1 A1 1 90 100 117 VKAAVDLSHF C F C F C F C F C F C F C F C F C F C F	30 RACE ADORES ADORES ADORES ADORES ADORES ADORES ADORES ADORES ADORES ADDRES A	SQR SQR KK	40 50 V QD AVSQ D C V QD AVSQ D C V = 120 RQD1L DLWIYLITQGY E V N 1 200 210 DSKLA FHWVARELUP K H H S LR R Q	ALTSSHTAAT AA SP N 140 EVFNRC Y D Y D FY - ::	10 140 150 160	EE-EEVGFPV E SD 160 CYKLVPVEP F HS

Dediter. Driede

	A	•				<u>**</u>	<u> </u>	
	A LAVbru	vs.	HTLV-3	ARV-2	LAVeli	LAV mat	L AVeli vs.	LAVmal
	5		512	502	500	505		505
Herst Heart George	GAG	,	0.8	3.4	9.8	12.0		10.8
the state of the s			201	1003	1370	1002		1002
	POL		1.3	3.1	5.5	7.7		8.4
dian half		입	950	055	053	059 13/11		13.11
		Total	1.4	13.0	653 22.14	21.7		19.8
Errill Erril	<u> </u>	0	507	505	504	509 13/10		509
	ENV	OMP	1.6	14.3	25.3	26.4		23.6
		T	349	350	349	350		350
		TMP	1.1	11.2	13.8	14.9		14.3

<	l'AVbru	orf	L.		central region	a	regior	_		
ζ	VS.	5	-	ort Q	0	0	A.	0	orf S	
	HTLV-3	206	1.5	192	0		pu	80 0,0	2.5	
	ARV-2	210	12.6	192	10.0	97	9.4	81	15.0	
		20ê	19.4	192	10.4	96	11.5	80	27.5	· · ·
	LAV mal	209	27.0	192	12.6	96	10.4	0 / 0	23.8	
8	L AVeli vs.									
	LAVmal	209	22.5	192	12.0	96 0 ⁄ 0	6.3	0,0	11.3	
										1



PIG. 5

-5 -5 -5 -5	0 0 0 0 0 0 0 0 0 0	
170	6 N F 1 0 S R P 1 P P F 1 0 S R P 1 P P F 1 C S R P C P P P F 1 P P F 1 P P F 1 P P F 1 P P F 1 P P F 1 P P F 1 P P F 1 P P F 1 P P P F 1 P P P F 1 P P P F 1 P P P P	N
GAG a uv uv	D LAY. 24. LAY. 24. LAY. 24.	C (4X,550 487 2 487 2 (4X,550)

COCLECT CACTOOC

LAV.anu LAV.nut LAV.nut LAV.nut LAV.nut LAV.nut LAV.nut LAV.nut	0 H 1 W 1 W 2 G 1 W 1 C 1 W 1 C 1 W 1 C 1 C C C C C C C
. — — — — — — — — — — — — — — — — — — —	4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

h w.m. LV.ML PYN ?

147.111

LAV. HAE

→F	}					=																																					:			
G	GI	C	T	C	T	CI	T	G	T	1	١G	A	C	CA	G	G	T(G	Λ	G	C	CC	G	G	G /	A G	C	T	CI	CC	TG	G	CI	A	G C	A	A	GG	A.	A (C C	C	A (CI	G	
							•								•								•			F	₹ &	4	> [J 5	•							•							•	
Ç,	TI	A	A	G	C	CI	C	A.	A'	T	۱A	A	G	C I	T	G	C	1	T	G	A (G 7	CG	C	C:	rc	A	A	GC	A	G 1	G'	TG	T	GC	; C	C	A I	C.	T(jŢ	T	G :	ΓG	Ī	
															•								•						1	0	0							•			ι	US	3	4	۲۰	
G.	A C	1	: C	T	G	G 1	A	Α	C,	T	A G	Α	G.	A?	r C	C	C:	0	A	G	A	C	A	C	T	C1	A7	G	A C	G	G1	G	TA	A.	AA	A	T	CI	C	T	A G	C	A(T	þ	
			_		-																		•								•							•							•	
G	cc	; (: c	С	G	A.A	C	Α	G	G	GÁ	C	T	T	ΓA	A	A	3]	G:	A	Ă.	A C	T	A	A	CA	G	G	GA	C	T	G	AA	A	GC	G	G	AA	G'	T	C	C	AC	3 A	G	
Ť	-		-	Ī	٠.				_						0 0																•							•								
A	A C	;]	T	С	T	C I	C	G	A	C	GC	: A	G	G Z	A C	T	C	3 (; C	T	T	G	I	G	A	GO	T	G	CA	LC.	A C	A	GC	A.	AG	A	G	GC	G	A(34	G	C	G G	C	
-	_		_	Ī	_			Ī		-	_				•																•							٠,		A	c		•	3 0	0	
							_																														Ī	Мe	t	G.	Гy	A	1 a	a A	r	8
G	A C	: 1	G	G	T	G Å	\ G	T	A	C	GC	c	A	A'	ΓI	T	T	ro	; A	C	T.	A (3 0	G	G	AC	G	C	T A	\G	A I	\C	GA	G	ΔG	A	d	AI	G.	G	jī	G	C	3 A	G	_
_				Ī	_		•									,															•						•	•							•	
	A 1	ءا	ı S	۾	T .	V٤	. i	L	e	u :	Sε	. T	G	1,	γĠ	1	y i	L	, s	L	e	u A	\ s	P	A	1 &	ıT	r	ρÓ	; 1	u I	y	s I	1	e A	ľ	8	Le	u	A:	r g	P	r	o G	1	y
Δ	ი (: 0	. T	Ċ	Ā	G 1	ΓA	T	Ť.	A	A G	C	G	G	GG	G	Á	A A	A	T	T.	A (; E	Ţ	G	C #	T	G	ĞC	GA	G A	۱Ă.	AA	T	TO	G	G:	TI	'A	A(ĴĞ	C	CA	A G	G	•
•	•	•		Ĭ	4.	٠.		_	-																	_			4	40	0			:				•					_	_	•	
	<u>ر</u> 1	١.	. 1	•	_	т.	•	T	17		т.	, -	. Δ	.	- - T	۵	,, ì	r. s	, e	ĸ	i	s I	. e	11	♥.	a 1	I T	٠,	n A	11	a 9	i e	T A	. T	9 (: 1	น	i.e	11	G '	l u	A	Tſ	e P	h	e
_	G .	ני	, <u>1</u>	y	·	נ ש מ	/ S		y	3 ·	<u>.</u>	T	· A	<u> </u>	5 ±	, C	د	ر ک ۵ ۸	Δ	~	Δ,	T '	ר ז		د	T 4	T		C (GC	 A /	. C	C A	. C	C (G	CT	<u>۔</u>	c /	 4 A	Ā	د د	a T	· T	_
G	G	s E	A	.A	. .	W.	\ A	A	A	n.	T 1.		A	G 1	n (•	G /		•	. •	•	•	• •		•		•		•	, ,		•	•		•	- 41	. •			•	242		٠.	•	-	
	A 1	١.	. 7	_		۸.	•	. 5	_	_ (c 1	١,,	, T	_	• • • •	۰ ـ		1 ي		т	h	- (; ;	v	G	١,	١G	1	v (Су	9 (: 1	n G	1	n I	1	el	H.	t	G	1 บ	G	11	nΙ	ė	81
c	A.	Lä	1 1	e	u.	A i	5 II	יר	-	T	ניט מים	. y	יר	T	TI	. T	Δí	. A	Δ	Ā	č	Δ	30	. J : A	<u>.</u>	À	. C	Ġ	ΑΊ	CG	T	. A	A C	Ā	Ä	Ī	À	A 1	G	G.	AA	C	Ā	GC	T	_
C	6 (u E	3 C	. 1		N I		. •		Τ,	9	, (, .		0 0			.	•••	•••	•	•••			•					•	•	_		•						-			•••	-	-	
	۲.	١.			_	ጥነ	•		_		T.		· T					٠.		c	1	(: 1		т	1 4	. 7	. 👽		S e	- - 1	۵.	., T	٠,	T #	۱.	n'	T h		₹.	a 1		1.	a T	'h	~
	G .	LI	12) e	A	TI	ת מית		T	u.	ر ب <u>ا</u>	7. 3 . C	. A	<u></u>	A C		J .	T (c	Δ	Δ		Δ	•	T'	r a	A	AT	C	41	T	AΊ	À	T	À	T	A C	A	Ġ	TΑ	G	Ĉ.	ΛA	c	-
A	U Z	1 4	4 1	. •	A	W,	· 1		+	.	u s	20	A			,	a	٠,	, ,	. •	•		ء د		••	•	• 41	•••	•••			•		•	••	•••			•••	•	• • •		1	6 0	0	
			. 7	•		۲.	•	, T7	, _	1	11 3	: .	. ~	٦.	_ /	, \	_	7 1	م ا	. A		~ ,	, 7 -	, 1	т.	.,	. A		~ 1	Γh	• •1		e ()	: 1	4	1	9,	1. a		.	e r	٠T.	v	e T	1	
_	L(e 1	 7 T	y	T	T	y 5	· ~	a	i A	n i	L 3		A	M F		Š	1 1 1 1	. e		a A	ያ ተሰ	77		<u></u>	y: Ai	. C	. s	C 1	A C	C A	. J	ເດ	1	AC			C 1	Δ.	G.	ΔT	· 🚡	A.	A A	Ť	•
U	U.	1 '	ر ي		. 1	Τ,	.		1	n	U Z			Λ.	n.		G ,				•	•			-				•		٠.	•••	•	•••	•••				•••	•	-	•	•••		· -	
	_	١.	. ~	٠,		.	1 -	· · ~		_	A .		. 7		_ (·	_	Α.	- ~	c	1	- 1		, .	т	h ,	- C	: 3	~ (31	n /	1	. .	1	• 1	۱ 1	•	A 1	2	G.	1 1	G	1 1	n A	. 1	
	6	TI	26	, 1	.u	1	1 E	96		n.	A :	3 F	1 Ju	y	5 i	. ~		r (. 8		Ā	<u></u>	ر ت	,	Ā			A	ci	CA	C.C	: .	A C	Ċ	40	: .	T	ה ה	. A	c	4 C	. C	A (c c	c	_
A	G	V.	ی و	, P.	A	A	1 8	٠.	A	A	M !		A	A	G S	7.0	U,	a (, ,	, ,	n	n,	74.F	70	•	C Z		'n	3	7 N	o (, .	AU	. •	av	, ,	•		44	٠.	20		•	-		
							. •	٠ _			•			_	_ 6			٠.			_	_,		. 1	c		· _	٠,	- <i>i</i>	As	٠,	P.,	_ t		۸ ۱	7 1	_	v.	. 1	c	1 -			n 4	. 1	•
	A	1	a	1 1	. a	A	la	1 1	Ω.	r	T.	y s	3 A	. S	n) e	I	21	: I		e	I T	V &	5 T	3	e :	T (, <u>.</u>	M A	 	T.	r y	Lt		C	. 4	. E	. T	C	C		. A	A 1	и <i>л</i> ТС		•
A	G	C	A (; (T	G	CC	A	C	P.	A	R.P	۱A	A	U I	ى 1	U.	v	, (, A	G	Τ,	ر ی		A	.	1 C	A	n.	۸A		LA			U Z		. 4	.	, G	U 1	241	144		+ 4		
	_	_		_		_	٠	•							_ '	•	_			. +	,	_	٠,	•	70		_ A		_,	ተ ኤ	• -1		1		- 4	A 1		• T-		v	. 1	7	17	. T	• } a	1
	G	1	ם (;]	y	G	11	1 .	i e	t	1.	16	H	1	5 (3 L	Π.	A.	la	, '	T	e	D (2 I	r	r	7 A	II	Š	I h	7	r T	ur Cit		T	. C	. a	T.(. C	c'	T.C	. <u>.</u>) A	4 C	- T	•
A	C.	A.	A (; (G	C.	A	1.	I	G	A.	I F	3 C				G	6		, A	. 1	A	Τ,	<i>-</i> ^	·	.	1 4		61	A C	1.		9.5	ın		3 C	, "	+ 4		.	10	, .	A	nv		
	_			_·-	<u>-</u>	_	_ '	• _				_	_		0(_		_			•••	•	.	,		٠	_ 1	٠.	•	n L	_ (, _	_	A 1		•		c		- ~	٦.	r	•	.,
	I	1	es	3)	Lu	G	Į١	u I	. y	5	A	li	3 F	'n	6	> €	I	Y :	rc	96	. 1	u	v :	1 1 1		T (ez	I I	01	i.e	C 1	r II TT	C1) E				Ti	; u	3) T	€ Y		A	c) L	J
å	ιA	T				G	A	A.A	\ A	G	G	C:	ГТ	. 1	C	Ą	C	C	ء ب	٠. ن	A	.P.	.	ט ג	A	1.4	A	<i>-</i> ر	U.	ΑT	G.	ΤŢ	U,			٠,	, A	1 1	. ^	1	C Z	. G	D.	9 (, G) N	
		_		•••	, To	_		•		•						•		.,			•	_		•		_	_ •			17 _	•	ر ا		- 1	1	: :	: _	• C 1			١,			•		
	A	1	a :	Γl	ı	P	r	0(; l	n.	A	s į	bΤ	, e	u.	A. S	n	11	e i	1 2	ıe		L (eu	A	5	u_1	1 1	e	V a	T	د د 1 م	y	 . T	. y :	C 1	. C	c i		n C	CI			TI	1 C	, -
(G	С	C	ሉ (CC	: C	C	A (3.5	A	G.	Α.	ΓΊ	T	A	A.F	. 1	A.	1 (j F	1	G	C	I G	Ā	A	Ų į	7	A	GI	Ι.	ن ب	A	, 6	A	<i>د</i> ک	٠.	Ų į	10	G	U 2	, 0	·	1.4		
						_		•					_			•				_			_	•							•	Φ	_		_			* *7 .	. 1	ш	: .	- T		_ 1	•	1
	G	1	n!	M	e t	L	e 1	u I	ر با	S	A	S	ρŢ	h	I	11	. е	A	SI	n (1	. u	G	1 .	ı Eı	. 1	a /	9 T	. a.	A s	, D	TI	P	4 S	p.	A 1	g	A 1	2 L	U	1:	3 F T C	·C	6 C	ነ ፈ ጉጥ	, 1
(C	A	Α.	A'	T C	T	T.	A A	A A	A	G	A.	ΓΑ	C	C	K]	C	A.	Α	T C	A	G	, ی	r. F	کالا	U	1.() و		G A		16	. (, (J A		. (, 6	G.	. A	. .	e. 1		, ب	۲. (3 L	
					_		_	•	_		_	_		_		•		_		_		1	٠.	•					-	00	•		٠.	~ •		c		•		. т	1	٠,	. 1	- 1	• •	4 .1
	Н	i	S	A	1 a	3 G	1	y i	P 1	0	I	1	e I	? I	0	P 1	0	G	1	y (,]	n	řĬ	e 1	A	I	å (G]	u	Pı	0	A I	8	[و م	y	2 (e I	Ā	s P	1	16	e A A c	, V	A /	3 C	, :
	A C	A	T	G	C &	٦G	G	G	C(: 1	Α.	T	T(C	C	C(J.A.	G	G	C (į	C	A	1 ('n	Ğ	A	J.	A	CC	·Λ	A	A	يا و	A	V.	o I	64	A U	٨	1 /	A (, ,	w.	, 0	,

The	and the	rreuGlu	GluGlnI	leGlyTr	pMetThrS	erAsnProB	rolleProVal
							CTATCCCAGT
3	•	110	₹.	•	•	3	•
GlyAs1	plleTy	rLysArg	TrpIleI	leLeuGl	yLeuAsnLy	ysIleVal <i>A</i>	rgMetTyrSer
GGGAGA	CATCTA	TAAAAGA	TGGATAA	TCCTGGG	ATTAAATAA	AAATAGTAA	GAATGTATAG
			•	•	•		1200
ProVa	lSerIl	eLeuAsp	IleArgG	InGlyPr	oLysGluPr	CoPheArgA	spTyrValAsp
CCCTGT	CAGCAT	TTTGGAC	ATAAGAU	AAGGGCC	AAAGGAAC	IIIIAGAG	ACTATGTAGA
A = a Ph	• • Pholy	eTh-lau	• A==A1=G		• Th=GlnG1	• 4 • • • • • • • • • • • • • • • • • •	.suTrpHetThr
							ATTGGATGAC
1				•	1300		
GluTh	LeuLe	uValGln.	AsnAlaA	snProAs	pCysLysTi	arlleLeuL	ysAlaLeuGly
							AAGCATTAGG
	•		•	•	•	•	•
							lyProSerHis
ACCAGG	GGCTAC.			TGACAGO	ATGCCAGG	GAGTGGGAG	GACCCAGTCA
	•	140	•	•	•	•	•
LysAl	aArgVa	lLeuAla	GluAlaM	etSerGl	nAlaThrAs	nSerThrA	laAlaIleMet
TAAAGC	AAGAGT	TTTGGCT	GAGGCAA	TGAGCCA	AG CAACAA	ATTCAACTG	CTGCCATAAT 1500
V C1.	- 4 1		• . !aC !C	•	•	• • • • • • • • • • • • • • • • • • • •	ysGlyLysGlu
CATCCAC	GYCYCC	YASHFRE. TAATTTT	TARCECC	TOTABAC	AATTAAGTO	TTTCAACT	GTGGCAAAGA
GRIGON	JAUAGG	1001114	ANG GG C C	nonnano •			• •
GluHi	s Leu A l	a Aro Asn	CvsArgA	laProAr	eLvsLvsG1	LvCvsTrpL	ysCysGlyLys
							AATGTGGGAA
	•		•	•	1600	POL .	•
	•		•	•	1600	PhePheAr	gGluAsnLeu
				hrGluAr	1600 gGlnAlaAs	PhePheAr nPheLeuG	lyLysIleTrp
				hrGluAr	1600 gGlnAlaAs	PhePheAr nPheLeuG	
GGAAGG	ACACCA.	AATGAAA	GACTGCA	hrGluAr CTGAGAG	1600 gGlnAlaAs ACAGGCTAA	PhePheAr inPheLeuG ITTTTTAG	lyLysIleTrp GGAAAATTTG
GGAAGG. AlaPhe	ACACCA. ProGln	AATGAAA Glylysa	GACTGCA • laArgGl	hrGluAr CTGAGAG	1600 gGlnAlaAs ACAGGCTAA SerGluGln	PhePheAr mPheLeuG TTTTTTAG nThrArgAl	TyLysIleTrp GGAAAATTTG aAsnSerPro
GGAAGGAAlaPhe ProSe	ACACCA ProGln rHisLy	AATGAAA GlyLysA sGlyArg	GACTGCA • laArgGl ProGlyA	hrGluAr CTGAGAG uPhePro	1600 gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr	PhePheAr mPheLeuG TTTTTTAG ThrArgAl gProGluP	iyLysileTrp GGAAAATTTG aAsnSerPro roThrAlaPro-
GGAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	ACACCA ProGln rHisLy	AATGAAA GlyLysA sGlyArg	GACTGCA • laArgGl ProGlyA	hrGluAr CTGAGAG uPhePro	1600 gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr	PhePheAr mPheLeuG TTTTTTAG ThrArgAl gProGluP	TyLysIleTrp GGAAAATTTG aAsnSerPro
AlaPhe ProSe GCCTTC	ACACCA ProGln rHisLy CCACAA	AATGAAA GlyLysA sGlyArg GGGAAGG 170	GACTGCA . laArgGl ProGlyA CCAGGGA 0	hrGluAr CTGAGAG uPhePro snPheLe	gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG	PhePheAr mPheLeuG TTTTTTAG ThrArgAl gProGluP GACCAGAGO	TyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro- CAACAGCCCC
AlaPhe ProSe GCCTTC	ACACCA ProGlu rHisLy CCACAA ArgGlu	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl	hrGluAr CTGAGAG uPhePro snPheLe ATTTCCT	1600 gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG	PhePheAr mPheLeuG TTTTTTAG ThrArgAl gProGluP GACCAGAGC	TyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro- CAACAGCCCC rGlyAlaGlu
AlaPhe ProSe GCCTTC ThrSer ProAl	ACACCA ProGln rHisLy CCACAA ArgGlu aGluSe	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG	hrGluAr CTGAGAG uPhePro snPheLe ATTTCCT yGlyAsp	1600 gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe	PhePheAr mPheLeuG TTTTTTAG ThrArgAl gProGluP GACCAGAGC uSerGluTh	TyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro- CAACAGCCCC
AlaPhe Prose GCCTTC ThrSer ProAl ACCAGC	ACACCA ProGlu rHisLy CCACAA ArgGlu aGluSe AGAGAG	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG	chrGluAr CTGAGAG uPhePro snPheLe ATTTCCT yGlyAsp luGluIl	gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe AAAACCCTG	PhePheAr mPheLeug ThrArgAl gProGlup ACCAGAGO SerGluTh rGlnLysG	TyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro- CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800
AlaPhe ProSe GCCTTC ThrSer ProAl ACCAGC	ACACCA ProGlu rHisLy CCACAA ArgGlu aGluSe AGAGAG	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG heSerPh	CTGAGAG UPhePro snPheLe ATTTCCT yGlyAsp cluGluIl GAGGAGAT	gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAn TCAGAGCAG LysThrLeu eLysProSe AAAACCCTG	PhePheAr mPheLeuG TTTTTTAG ThrArgAl gProGluP GACCAGAGC SerGluTh erGlnLysG TCAGAAAC	TyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro- CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal
AlaPhe ProSe GCCTTC ThrSer ProAl ACCAGC	ACACCA ProGlu rHisly CCACAA ArgGlu aGluSe AGAGAG GlyIle	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG ValSerP uTyrPro	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG heSerPh LeuAlaS	CTGAGAG UPhePro SnPheLe ATTTCCT UG1yAsp SluG1uI1 AGGAGAT	gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe AAAACCCTG	PhePheAr mPheLeuG TTTTTTAG ThrArgAl gProGluP GACCAGAGC SerGluTh erGlnLysG CTCAGAAAC TrpGlnAr neGlyAsnA	IyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro- CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal spGlnLeuSer
AlaPhe Prose GCCTTC ThrSer ProAl ACCAGC ArgCTA AGA	ACACCA ProGlu rHisly CCACAA ArgGlu aGluSe AGAGAG GlyIle	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG ValSerP uTyrPro	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG heSerPh LeuAlaS	CTGAGAG UPhePro SnPheLe ATTTCCT UG1yAsp SluG1uI1 AGGAGAT	gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe AAAACCCTG	PhePheAr mPheLeuG TTTTTTAG ThrArgAl gProGluP GACCAGAGC SerGluTh erGlnLysG CTCAGAAAC TrpGlnAr neGlyAsnA	TyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro- CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal
AlaPhe Prose GCCTTC ThrSer ProAl ACCAGC ArgCA AsgA AGA	ACACCA ProGlu rHisLy CCACAA ArgGlu aGluSe AGAGAG GlyIle GGAATT	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG ValSerP uTyrPro GTATCCT	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG heSerPh LeuAlaS TTAGCTT	chrGluAr CTGAGAG uPhePro snPheLe ATTTCCT yGlyAsp luGluIl GAGGAGAT eProGln cerLeuLy	gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe AAAACCCTG	PhePheAr mPheLeug TTTTTTAG ThrArgAl gProGluP GACCAGAGO SerGluTh erGlnLysG TCAGAAAO TTPGlnAr eGlyAsnA	lyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal spGlnLeuSer ACCAGTTGTC
AlaPhe ProSe GCCTTC ThrSer ProAl ACCAGC ArgC	ACACCA ProGlu rHisLy CCACAA ArgGlu aGluSe AGAGAG GlyIle GGAATT	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG ValSerP uTyrPro GTATCCT	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG heSerPh LeuAlaS TTAGCTT	chrGluAr CTGAGAG uPhePro snPheLe ATTTCCT yGlyAsp luGluIl GAGGAGAT eProGln cerLeuLy	gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe AAAACCCTG	PhePheAr mPheLeug TTTTTTAG ThrArgAl gProGluP GACCAGAGO SerGluTh erGlnLysG TCAGAAAO TTPGlnAr eGlyAsnA	IyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro- CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal spGlnLeuSer
AlaPhe ProSe GCCTTC ThrSer ProAl ACCAGC ArgGIA AGA ThrVal Gln	ProGlurHisLyCCACAAAArgGluSeAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG ValSerP uTyrPro GTATCCT GlyGlyG	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG heSerPh LeuAlaS TTAGCTT lnLeuLy	CTGAGAG UPhePro SnPheLe ATTTCCT UGlyAsp CluGluIl AGGAGAT CEProGln CETLEUL CCCTCAA	gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe AAAACCCTC IleThrLeu sSerLeuPh ATCACTCTT	PhePheAr mPheLeug ThrArgAl gProGluP GACCAGAGO TSETGIUTH CTGINLYSG CTCAGAAAO TTPGINAT MEGIYASNA TTGGCAACG	TyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro- CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal spGlnLeuSer ACCAGTTGTC aAspAspThr
AlaPhe ProSe GCCTTC ThrSer ProAl ACCAGC ArgGIA AGA ThrVal Gln	ProGlurHisLyCCACAAAArgGluSeAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG ValSerP uTyrPro GTATCCT GlyGlyG	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG heSerPh LeuAlaS TTAGCTT lnLeuLy	CTGAGAG UPhePro SnPheLe ATTTCCT UGlyAsp CluGluIl AGGAGAT CEProGln CETLEUL CCCTCAA	gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAn TCAGAGCAG LysThrLeu eLysProSe AAAACCCTC IleThrLeu sSerLeuPh ATCACTCTT LeuLeuAsp	PhePheAr mPheLeug ThrArgAl gProGluP GACCAGAGO TSETGIUTH CTGINLYSG CTCAGAAAO TTPGINAT MEGIYASNA TTGGCAACG	lyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal spGlnLeuSer ACCAGTTGTC
AlaPhe ProSe GCCTTC ThrSer ProAl ACCAGC ArgGA AsgA ThrVal Gln ACAGTA	ACACCA ProGlu rHisly CCACAA ArgGlu aGluSe AGAGAG GlyIle GGAATT ArgVal AGAGTA	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG ValSerP uTyrPro GTATCCT GlyGlyG GGAGGAC	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG heSerPh LeuAlaS TTAGCTT lnLeuLy AGCTAAA	CTGAGAG UPhePro SnPheLe ATTTCCT YG1yAsp SluGluII AGGAGAT CEPTOGIN CETLEULY CCCTCAA	1600 gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe AAAACCCTC IleThrLeu sSerLeuPh ATCACTCTT LeuLeuAs; CTATTAGAG	PhePheAr mPheLeug TTTTTTAG ThrArgAl gProGluP GACCAGAGO SerGluTh erGlnLysG TCAGAAAO TTPGlnAr neGlyAsnA TTGGCAACG ThrGlyAl	IyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro- CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal spGlnLeuSer ACCAGTTGTC aAspAspThr AGGATGATACA
AlaPhe ProSe GCCTTC ThrSer ProAl ACCAGC ArgGA ArgGA ArgGA ThrVal GAG CAGTA ValLeu	ProGlurHisLy CCACAA ArgGluSe AGAGAG GlyIle GAATT ArgVal AGAGTA GluGlu	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG ValSerP uTyrPro GTATCCT GlyGlyG GGAGGAC IleAsnL	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG heSerPh LeuAlaS TTAGCTT lnLeuLy AGCTAAA euProGl	CTGAGAG UPhePro SnPheLe ATTTCCT YG1yAsp SluGluII AGGAGAT GEPTOGIN SCCCTCAA YSG1uAla AAGAAGCT	1600 gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe AAAACCCTG IleThrLeu sSerLeuPh ATCACTCTT LeuLeuAsp	PhePheAr mPheLeug aThrArgAl gProGluP GACCAGAGO aSerGluTh erGlnLysG TCAGAAAO aTrpGlnAr erGlyAl CACAGGAGO ThrGlyAl	TyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal spGlnLeuSer ACCAGTTGTC aAspAspThr AGATGATACA yGlyIleGly
AlaPhe ProSe GCCTTC ThrSer ProAl ACCAGC ArgGA AST AGA ThrVal Gln ACAGTA ValLeu GTATTA	ACACCA ProGlu rHisLy CCACAA ArgGlu aGluSe AGAGAG GlyIle GAATT ArgVal AGAGTA GluGlu GAAGAA	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG ValSerP uTyrPro GTATCCT GlyGlyG GGAGGAC IleAsnL ATAAATT	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG . heSerPh LeuAlaS TTAGCTT . lnLeuLy AGCTAAA . euProGl	ChrGluAr CTGAGAG uPhePro snPheLe ATTTCCT yGlyAsp SluGluII GAGGAGAT GETLEULY CCCTCAA ySGluAla AGGAAGCT GAAAATGG	1600 gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe AAAACCCTC IleThrLeu sSerLeuPh ATCACTCTI LeuLeuAs 1900 LysProLys AAAACCAAAA	PhePheAr mPheLeug aThrArgAl gProGlup GACCAGAGO aSerGluTh erGlnLysG TCAGAAAO aTrpGlnAr aCGlyAsnA aTrpGlnAr aCGLyAsnA CACAGGAGO ATTGCAACG ATTGCAACG	TyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal spGlnLeuSer ACCAGTTGTC aAspAspThr AGGATGATACA yGlyIleGly GGGAATTGGA
GGAAGG AlaPhe ProSe GCCTTC ThrSer ProAl ACCAGC ArgGA ASA AGA ThrVal Gln ACAGTA ValLeu GTATTA GlyPhe	ACACCA ProGlu rHisLy CCACAA ArgGlu aGluSe AGAGAG GlyIle GAATT ArgVal AGAGTA GluGlu GAAGAA IleLys	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG ValSerP uTyrPro GTATCCT GlyGlyG GGAGGAC IleAsnL ATAAATT	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG heSerPh LeuAlaS TTAGCTT lnLeuLy AGCTAAA euProGl TGCCAGG	ChrGluAr CTGAGAG uPhePro snPheLe ATTTCCT yGlyAsp cluGluII GAGGAGAT ceProGln cerLeuLy CCCTCAA ySGluAla AAGAAGCT tyLysTrp GAAAATGG	1600 gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe AAAACCCTO IleThrLeu sSerLeuPh ATCACTCTT LeuLeuAsp CTATTAGAG 1900 LysProLys AAACCAAAA	PhePheAr mPheLeug ThrArgAl gProGlup aCCAGAGO aSerGluTh erGlnLysG TCAGAAAO aTrpGlnAr aCGlyAsnA aTrgGCAACO aThrGlyAl cACAGGAGO sMetIleGl AATGATAGO	TyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro- CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal spGlnLeuSer ACCAGTTGTC aAspAspThr AGATGATACA yGlyIleGly GGGAATTGGA
GGAAGG AlaPhe ProSe GCCTTC ThrSer ProAl ACCAGC ArgGA ASA AGA ThrVal Gln ACAGTA ValLeu GTATTA GlyPhe	ACACCA ProGlu rHisLy CCACAA ArgGlu aGluSe AGAGAG GlyIle GAATT ArgVal AGAGTA GluGlu GAAGAA IleLys	AATGAAA GlyLysA sGlyArg GGGAAGG 170 LeuArgV rPheGly CTTCGGG ValSerP uTyrPro GTATCCT GlyGlyG GGAGGAC IleAsnL ATAAATT	GACTGCA laArgGl ProGlyA CCAGGGA 0 alTrpGl PheGlyG TTTGGGG heSerPh LeuAlaS TTAGCTT lnLeuLy AGCTAAA euProGl TGCCAGG	ChrGluAr CTGAGAG uPhePro snPheLe ATTTCCT yGlyAsp cluGluII GAGGAGAT ceProGln cerLeuLy CCCTCAA ySGluAla AAGAAGCT tyLysTrp GAAAATGG	1600 gGlnAlaAs ACAGGCTAA SerGluGln uGlnSerAr TCAGAGCAG LysThrLeu eLysProSe AAAACCCTO IleThrLeu sSerLeuPh ATCACTCTT LeuLeuAsp CTATTAGAG 1900 LysProLys AAACCAAAA	PhePheAr mPheLeug ThrArgAl gProGlup aCCAGAGO aSerGluTh erGlnLysG TCAGAAAO aTrpGlnAr aCGlyAsnA aTrgGCAACO aThrGlyAl cACAGGAGO sMetIleGl AATGATAGO	TyLysIleTrp GGAAAATTTG aAsnSerPro roThrAlaPro CAACAGCCCC rGlyAlaGlu lnGluGlnLys AGGAGCAGAA 1800 gProValVal spGlnLeuSer ACCAGTTGTC aAspAspThr AGGATGATACA yGlyIleGly GGGAATTGGA

llegizzne!	ITeLeuVa	LIGLYPTOTHEP	rovalAsnIl	elleGlyArgAsnHetLeuTh
ATAGGTACAA	LTATIGGI	AGGACCTACAC	CTGTCAACAT	AATTGGACGAAATATGTTGAC
•	•	• '	•	• • 210
GlnIleGlyC	ysThrLe	uAsnPheProI	leSerProI1	eGluThrValProValLysLe
CAGATTGGTT	GTACTT	AAATTTTCCAA	TAGTCCTAT	TGAGACTGTACCAGTAAAATT
•		•	•	
LvsProGlvM	letAspG1	vProAreValL	vsGlnTrnPr	oLeuThrGluGluLysIleLy
AAGCCAGGA	TEGATEG	CCCAAGGGTTA	AACAATCCCC	ATTGACAGAAGAAAAATAAA
·	11007100	CCCAAGGIIA	22	
• • • • • • • • • • • • • • • • • • • •	. 1 7 1 C			
CCATTALCAC	, id itech	SLYSASPHETG.	Innaggrand	yLysIleLeuLysIleGlyPr
GUALTAACAG	AAATIIG	TAAAGATATGG	AAAGGAAGG	AAAAATTTTAAAAATTGGGCC
C1A == D== m		•	•	•
GluasnProl	yrAsnTh	TPTOValPheA	lalleLysLy	sLysAspSerThrLysTrpAr
GRAAATCCAT	ACAATAC	TCCAGTATTTG	CCATAAAGAA	AAAAGACAGCACTAAATGGAG
		300	•	•
LysLeuValA	snPheAr	gGluLeuAsnLy	sArgThrG1:	nAspPheTrpGluValGlnLe
AAATTAGTGA	ATTTCAG	AGAGCTTAATAA	AAGAACTCAA	AGATTTTTGGGAAGTTCAATT.
•		•	•	. 240
GlyIleProH	isProA1	aGlyLeuLysLy	sLysLysSei	ValThrValLeuAspValG1
GGAATACCAC	ATCCTGC	TGGGTTGAAAA	GAAAAAATCA	AGTCACAGTATTGGATGTGGG
. •		•	•	•
AspAlaTyrP	heSerVa	1ProLeuAspG1	uAspPheArg	LysTyrThrAlaPheThrIl
GATGCATATT	TTTCAGT	CCCTTTAGATGA	AGATTTCAG	AAGTATACTGCATTCACTAT
•		•	. 250	00
ProSerIleA	snAsnG1	uThrProGlvI1	eArgTvrGlr	TyrAsnValLeuProGlnGl
CCCAGTATTA	ATAATGA	GACACCAGGGAT	TAGATATCAG	TACAATGTGCTACCACAGGG
		•	•	
TrpLvsGlvS	erProAl	allePheGlnSe	TSerWetThr	LysIleLeuGluProPheArg
TGGAAAGGAT	CACCAGO	AATATTCCAGAG	TAGCATGACA	AAAATCTTAGAACCCTTTAG
	2	600		MARKICITAGRACCCITIAGE
ThrivsAsrP			nTurketter	AspLeuTyrValGlySerAs
ACAAAAAATC	CAGAAAT	ACTCATATACCA	ATACATCCAT	GATTTGTATGTAGGGTCTGAT
onnanaa t	CAGRAAT.	AG ICA IA IACCA	AINCAIGGEI	
LauGlullaG	1.001-#4	• • A = • T b = T = • T 1	• • • • • • • • • • • • • • • • • • • •	• 2700 ArgGluHisLeuLeuLysTrp
TTACAAATAC	CACAACA	evidinimàs i i	GLUGIULEU	ArgGlumisLeuLeuLysTr
IINGAAAIAG	GACAACA	IAGAACAAAAI	AGAGGAACTA	AGAGAACATCTATTGAAATG
ClDb - Tb - T	5 - 7 4 -	- 1 . 1 . 11 . 0 .	•	•
CCATTTAGG	DIFTOAS	brasrasmise r	nLysGluPro	ProPheLeuTrpHetGlyTyr
GGATTIACCA	CACCAGA	CAAAAAGCATCA		CCATTTCTTTGGATGGGGTAT
			. 280	
GIULEBHISP	TOASPLY	sTrpThrValGl	nProlleGin	LeuProAspLysGluSerTrp
GAACTOCACC	CTGACAA.	ATGGACAGTGCA	GCCTATACAA	CTGCCAGACAAGGAAAGCTG
TV 11 1 1 1	4		•	•
INTVALASTA	splieGl	nLysLeuValGl	yLysLeuAsn	TrpAlaSerGlnIleTyrPro
AUTGTCAATG.			AAAACTAAAT	TGGGCAAGTCAGATTTATCCA
•	_	900	•	•
GlylleLysV	alLysG1	nLeuCysLysLe	uLeuArgCly	AlaLysAlaLeuThrAspIla
GGAATTAAAG	TAAAGÇA.	ATTATGTAAACT	CCTTAGGGGA	GCAAAAGCACTAACAGACATA
•		_	_	. 3000
		•	•	• 3000
ValProLeuT	hrAlaG1	uAlaGluLeuGl	uLeuAlaGlu	AsnArgGluIleLeuLysGlu AACAGGGAAATTCTAAAAGAA

ProvelEisGlyValTyrTyrAspProSerLysAspLeuIleAlaGluIleGlnLysGln CCAGT CATGGGGTATATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAG GlyGlaGlyGlnTrpThrTyrGlnIleTyrGlnGluGlnTyrLysAsnLeuLysThrGly GGGCAAGGTCAATGGACATATCAAATATACCAAGAGCAATATAAAAATCTGAAAACAGGG LysTyrAlaArgIleLysSerAlaHisThrAsnAspValLysGlnLeuThrGluAlaVal AAGTATGCAAGAATAAAGTCTGCCCACACTAATGATGTAAAACAATTAACAGAAGCAGTG 3200 GlnLysIleAlaGlnGluSerIleValIleTrpGlyLysThrProLysPheArgLeuPro CAAAAGATAGCCCAAGAAAGCATAGTAATATGGGGAAAAACTCCTAAATTTAGACTACCC IleGlnLysGluThrTrpGluAlaTrpTrpThrGluTyrTrpGlnAlaThrTrpIlePro ATACAAAAGAAACATGGGAGGCATGGTGGACAGAATATTGGCAAGCCACCTGGATCCCT GluTrpGluPheValAsnThrProProLeuValLysLeuTrpTyrGlnLeuGluThrGlu GAATGGGAGTTTGTCAATACTCCTCCCCTAGTAAAACTATGGTACCAGTTAGAAACAGAA ProlleValGlyAlaGluThrPheTyrValAspGlyAlaAlaAsnArgGluThrLysLys CCCATAGTAGGÁGCAGAAACTTTCTATGTAGATGGGGCAGCTAATAGÁGAAACTAAAAAG GlyLysAlaGlyTyrValThrAspArgGlyArgGlnLysValValSerLeuThrGluThr GGAAAAGCAGGATATGTTACTGACAGAGGAAGACAAAGGTTGTCTCCTTAACTGAAACA Thr Asn Glu Leu Glu Ala Ile His Leu Ala Leu Glu Asp Ser Gly Ser GluACAAATCAGAAGACTGAATTACAAGCAATCCACTTAGCTTTACAGGATTCAGGATCAGAA ValAsnIleValThrAspSerGlnTyrAlaLeuGlyIleIleGlnAlaGlnProAspLys GTAAACATAGTAACAGACTCACAGTATGCATTAGGGATTATTCAAGCACAACCAGATAAA SerGluSerGluIleValAsnGlnIleIleGluGlnLeuIleGlnLysAspLysValTyr AGTGAATCAGAGATTGTTAATCAAATAATAGAGCAATTAATACAGAAGGACAAGGTCTAC LeuSerTrpValProAlaHisLysGlyIleGlyGlyAsnGluGlnValAspLysLeuVal CTGTCATGGGTACCAGCACAAAGGGATTGGAGGAAATGAACAAGTAGATAAATTAGTC SerSerGlyIleArgLysValLeuPheLeuAspGlyIleAspLysAlaGlnGluGluHis 3800 GluLyaTyrHisSerAsnTrpArgAlaHetAlaSerAspPheAsnLeuProProIleVal GAAAATATCACAGCAATTGGAGAGCAATGGCTAGTGACTTTAATCTACCACCTATAGTA AlaLysGluIleValAlaSerCysAspLysCysGlnLeuLysGlyGluAlaMetHisGly GlnValAspCysSerProGlyIleTrpGlnLeuAspCysThrHisLeuGluGlyLysIle CAAGTAGACTGTAGTCCAGGGATATGGCAATTAGATTGCACACATCTAGAAGGAAAAATA IleIleValAlaValHisValAlaSerGlyTyrIleGluAlaGluValIleProAlaGlu ATCATAGTAGCAGTCCATGTAGCCAGTGGATATATAGAAGCAGAAGTTATCCCAGCAGAA ThrGlyGlnGluThrAlaTyrPheIleLeuLysLeuAlaGlyArgTrpProValLysVal ACAGGACAGGACAGCATACTTATACTAAAATTAGCAGGAAGATGGCCAGTAAAAGTA 4100

ValRendanGTySerAsnPheThrSerAlaAlaValLysAlaAlaCysTrpTrp CTACA PAGACAATGGCAGCAATTTCACCAGTGCTGCAGTTAAAGCAGCCTGTTGGTGG AlaAspIleLysGlnGluPheGlyIleProTyrAsnProGlnSerGlnGlyValValGlu GCAAATATCAAACAGGAATTTGGAATTCCCTACAACCCCCAAAGTCAAGGAGTAGTGGAA SerMetAsnLysGluLeuLysLysIleIleGlyGlnValArgGluGlnAlaGluEisLeu_ TCTATGAATAAGGAATTAAAGAAAATCATAGGGCAGGTAAGAGAGCAAGCTGAACACCTT LysThrAlaValGlnHetAlaValPheIleHisAsnPheLysArgLysGlyGlyIleGly AÁGA CAG CAG TA CAAA TGG CAG TG TT CAT TCA CAAT TT TAAAAG AAAA GGG GGG AT TGG G GlyTyrSerAlaGlyGluArgIleIleAspMetIleAlaThrAspIleGlnThrLysGlu GGGTACAGTGCAGGGGAAAGAATAATAGACATGATAGCAACAGACATACAAACTAAAGAA-LeuGlnLysGlnIleThrLysIleGlnAsnPheArgValTyrTyrArgAspAsnArgAsp TTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAACAGAGAC ProIleTrpLysGlyProAlaLysLeuLeuTrpLysGlyGluGlyAlaValValIleGln CCAATTTGGAAAGGACCAGCAAAACTACTCTGGAAAGGTGAAGGGGCAGTAGTAATACAG AspAsnSerAspIleLysValValProArgArgLysAlaLysIleIleArgAspTyrGly MetGlu GACAATAGTGATATAAAGGTAGTACCAAGAAGAAAAGCAAAAATCATTAGGGATTATGGA LysGlnMetAlaGlyAspAspCysValAlaGlyGlyGlnAspGluAsp AsnArgTrpGlnValMetIleValTrpGlnValAspArgHetArgIleArgThrTrpHis AAACAGATGGCAGGTGATGATTGTGTGGCAGGTGGACAGGATGAGGATTAGAACATGGCA SerLeuValLysHisHisHisHetTyrValSerLysLysAlaLysAsnTrpPheTyrArgHis CAGTTTAGTAAAACATCATATGTATGTCTCAAAGAAAGCTAAAAATTGGTTTTATAGACA His TyrGluSerArgHisProLysValSerSerGluValHisIleProLeuGlyAspAla TCACTATGAAAGCAGGCATCCAAAAGTAAGTTCAGAAGTACACATCCCACTAGGGGATGC ArgLeuValValArgThrTyrTrpGlyLeuGlnThrGlyGluLysAspTrpHisLeuGly TAGATTAGTAGTAAGAACATATTGGGGTCTGCAAACAGGAGAAAAAGACTGGCACTTGGG HisGlyValSerIleGluTrpArgGlnLysArgTyrSerThrGlnLeuAspProAspLeu TCATEGGTCTCCATAGAATGGAGGCAGAAAAGATATAGCACACAACTAGATCCTGACCT AlpepGlnLeuIleEisLeuTyrTyrPheAspCysPheSerGluSerAlaIleArgGln AG CAGA CCAA CTGATTCAT CTGTACTATTTTGATTGTTTTTCAGAATCTGCCATAAGACA AlaIleLeuGlyHisIleValSerProArgCysAspTyrGln&laGlyHisAsnLysVal AGCCATATTAGGACATATAGTTAGTCCTAGGTGTGATTATCAAGCAGGACATAACAAGGT 5000 GlySerLeuGlnTyrLeuAlaLeuThrAlaLeuIleAlaProLysLysThrArgProPro AGGAT CTTTA CAGTATTTG G CACTAACAG CATTAATAG CACCAAAAAAG A CAAG G C C & C C HetGluGlnAlaProAlaAspGlnGly LeuProSerValArgLysLeuThrGluAspArgTrpAsnLysProGlnGlnThrLysGly

TTTGCCTAGTGTTAGGAAGCTAACAGAAGATAGATGGAACAAGCCCCAGCAGACCAAGGG

LuProllisAsnGluTrpThrLeuGluLeuLcuGluGluLeuLysGlnGlu HisArgGlyS rHisThrMetAsnGlyHis CCACAGAGGGAGCCACACAATGAATGGACAHTAGAACTTTTAGAGGAGCTTAAGCAAGAA AlaValArgHisPheProArgIleTrpLeuEisSerLeuGlyGlnHisIleTyrGluThr GCTGTCAGACACTTTCCTAGGATATGGCTCCATAGTTTAGGACAACATATCTATGAAACT TyrGlyAspThrTrpGluGlyValGluAlaIleIleArgSerLeuGlnGlnLeuLeuPhe TATGGGGATACCTGGGAAGGAGTTGAAGCTATAATAAGAAGTCTGCAACAACTGCTGTTT 5300 IleHisPheArgIleGlyCysGlnHisSerArgIleGlyIleThrArgGlnArgArgAla ATTCATTTCAGAATTGGGTGTCAACATAGCAGAATAGGCATTACTCGACAGAGAAGAGCA ArgAsnGlySerSerArgSer MetAspProValAspProAsnLeuGluProTrpAsnHisProGlySerGlnProArg AGAALATGGATCCAGTAGATCCTAACTTAGAGCCCTGGAACCATCCAGGGAGTCAGCCTAG ThrProCysAsnLysCysTyrCysLysLysCysCysTyrHisCysGlnHetCysPheIle GACGCCTTGTAATAAGTGTTATTGTAAAAAGTGCTGCTATCATTGCCAAATGTGCTTCAT ThrLysGlyLeuGlyIleSerTyrGlyArgLysLysArgArgGlnArgArgArgProPro AACGAAAGGCTTAGGCATCTCCTATGGCAGGAAGAGCGGAGACAGCGACGAAGACCTCC GlnGlyAsnGlnAlaHisGlnAspProLeuProGluGln TCAGGGCAATCAGGCTCATCAAGATCCTCTACCAGAGCAG TAAGTAGTATATGTAATACA ACCTTTAGTGATATTAGCAATAGTAGCATTAGTAGTAACGCTAATAATAGCAATAGTTGT 5700 GTGGACCATAGTATTTATAGAAATTAGGAAAATAAGAAGACAAAGGAAAATAGACAGGTT MetArgValArgGluIleGlnArg **GATTGATAGAATAAGAGAAGAGCAGAAGATAGTGGCAATGAGAGTGAGGGAGATACAGA** 5800 AsnTyrGlnAsnTrpTrpArgTrpGlyMetHetLeuLeuGlyMetLeuMetThrCysSer GGAATTATCAAAACTGGTGGAGATGGGGCATGATGCTCCTTGGGATGTTGATGACCTGTA IleAlaGluAspLeuTrpValThrValTyrTyrGlyValProValTrpLysGluAlaThr GTATTG CACAAGATTTGTGGGTTACAGTTTATTATGGGGTACCTGTGTGGAAAGAAGCAA ThrTieLeuPheCysAlaSerAspAlaLysSerTyrGluThrGluValHisAsnIleTrp CCACTAGECTATTTTGTGCATCAGATGCTAAATCATATGAAACAGAAGTACATAACATCT AlaThrHisAlaCysValProThrAspProAsnProGlnGluIleGluLeuGluAsnVal ThrGluGlyPheAsnMetTrpLysAsnAsnMetValGluGlnMetHisGluAspIleIle TCACAGAAGGGTTTAACATGTGGAAAAATAACATGGTGGAGCAGATGCATGAGGATATAA 6100

SerLeuTrpAspGlnSerLeuLyaPr CysValLysLeuThrProLeuCysValThrLeuTCAGTTTACGGATCAAAGCCTAAAACCATGTGTAAAGCTAACCCCACTCTGTGTCACTT

AsnCysThrAsnValAsnGlyThrAlaValAsnGlyThrAsnAlaGlySerAsnArgThr
TAAACTGCACTAATGTGAATGGGACTGCTGTGAATGGGACTAATGCTGGGAGTAATAGGA
6200

AsnAlaGluLeuLysMetGluIleGlyGluValLysAsnCysSerPheAsnIleThrProCTAATGCAGAATTGAAAATGGAAATTGGAGAAGTGAAAAACTGCTCTTTCAATATAACCC

ValGlySerAspLysArgGlnGluTyrAlaThrPheTyrAsnLeuAspLeuValGlnIle CAGTAGGAAGTGATAAAAGGCAAGAATATGCAACTTTTTATAACCTTGATCTAGTACAAA

AspAspSerAspAsnSerSerTyrArgLeuIleAsnCysAsnThrSerValIleThrGln
TAGATGATAGTGATAATAGTTATAGGCTAATAAATTGTAATACCTCAGTAATTACAC

AlaCysProLysValThrPheAspProIleProIleHisTyrCysAlaProAlaGlyPheAGGCTTGTCCAAAGGTAACCTTTGATCCAATTCCCATACATTATTGTGCCCCAGCTGGTT

AlaIleLeuLysCysAsnAspLysLysPheAsnGlyThrGluIleCysLysAsnValSer TTGCAATTCTAAAGTGTAATGATAAGAAGTTCAATGGAACGGAAATATGTAAAAATGTCA 6500

Thr ValGlnCysThrHisGlyIleLysProValValSerThrGlnLeuLeuLeuAsnGlyGTACAGTACAATGTACACATGGAATTAAGCCAGTGGTGTCAACTCAACTGCTGTTAAATG

SerLeuAlaGluGluIleMetILeArgSerGluAsnLeuThrAspAsnThrLysAsnGCAGTCTAGCAGAAGAGAGATAATGATTAGATCTGAAAATCTCACAGACAATACTAAAA

IleIleValGlnLeuAsnGluThrValThrIleAsnCysThrArgProGlyAsnAsnThr ACATAATAGTACAGCTTAATGAAACTGTAACAATTAATTGTACAAGGCCTGGAAACAATA

ArgArgGlyIleEisPheGlyProGlyGlnAlaLeuTyrThrThrGlyIleValGlyAspCAAGAAGAGGGATACATTTCGGCCCAGGGCAAGCACTCTATACAACAGGGATAGTAGGAG

IleArgArgAlaTyrCysThrIleAsnGluThrGluTrpAspLysThrLeuGlnGlnVal ATATAAGAAGAGCATATTGTACTATTAATGAAACAGAATGGGATAAAACTTTACAACAGG . 6800

AlaValLysLeuGlySerLeuLeuAsuLysThrLysIleIlePheAsuSerSerSerGly.
TAGCTGTAAAACTAGGAAGCCTTCTTAACAAAACAAAAATAATTTTTAATTCATCCTCAG

GlyAspProGluIleThrThrHisSerPheAsnCysArgGlyGluPhePheTyrCysAsnGAGGGGACCCAGAAATTACAACACACAGTTTTAATTGTAGAGGGGAATTTTTCTACTGTA

Thr & LysLeuPheAsnSerThrTrpGlnAsnAsnGlyAlaArgLeuSerAsnSerThr ATACATCAAAACTGTTTAATAGTACATGGCAGAATAATGGTGCAAGACTAAGTAATAGCA

GluSerThrGlySerIleThrLeuProCysArgIleLysGlnIleIleAsnHetTrpGlnCAGAGTCAACTGGTAGTATCACACTCCCATGCAGAATAAAACAAATTATAAATATGTGGC

LysThrGlyLysAlaMetTyrAlaProProIleAlaGlyValIleAsnCysLeuSerAsn AGAAAACAGGAAAAGCTATGTATGCCCCTCCCATCGCAGGAGTCATCAACTGTTTATCAA

IleThrGlyLeuIleLeuThrArgAspGlyGlyAsuSerSerAspAsuSerAspAsuGluATATTACAGGGCTGATATTAACAAGAGATGGTGGAAATAGTAGTGACAATAGTGACAATG

7200

Thr Levarger GlyGlyGlyAspMetArgAspAsnTrpIleSerGluLeuTyrLysTyrAGACCTTAAGACCTGGAGGAGGAGATATGAGGGACAATTGGATAAGTGAATTATAAAT

GluArgGluLysArgAlaIleGlyLeuGlyAlaHetPheLeuGlyPheLeuGlyAlaAla TGGAAAGAGAAAAAGAGCAATAGGACTAGGAGCCATGTTCCTTGGGTTCTTGGGAGCAG

GlySerThrMetGlyAlaAlaSerLeuThrLeuThrValGlnAlaArgGlnLeuLeuSer CAGGAAGCACGATGGGCGCAGCGTCACTAACGCTGACGGTACAGGCCAGACAGTTACTGT

GlyIleValGlnGlnGlnAsnAsnLeuLeuArgAlaIleGluAlaGlnGlnEisLeuLeuCTGGTATAGTGCAACAGCAAAACAATTTGCTGAGGGCTATAGAGGCGCAACAGCATCTGT

GlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgValLeuAlaValGluArgTyrTGCAACTCACGGTCTGGGGCATTAAACAGCTCCAGGCAAGAGTCCTGGCTGTGGAAAGAT

LeuGlnAspGlnArgLeuLeuGlyMetTrpGlyCysSerGlyLysHisIleCysThrThr ACCTACAGGATCAACGGCTCCTAGGAATGTGGGGTTGCTCTGGAAAACACATTTGCACCA 7600

PheValProTrpAsnSerSerTrpSerAsnArgSerLeuAspAspIleTrpAsnAsnMetCATTTGTGCCTTGGAACTCTAGTTGGAGTAATAGATCTCTAGATGACATTTGGAATAATA

ThrTrpHetGlnTrpGluLysGluIleSerAsnTyrThrGlyIleIleTyrAsnLeuIleTGACCTGGATGCAGTGGGAAAAAGAAATTAGCAATTACACAGGCATAATATACAACTTAA

GluGluSerGlnIleGlnGlnGluLysAsnGluLysGluLeuLeuGluLeuAspLysTrp TTGAAGAATCGCAAATCCAGCAAGAAAAGAATGAAAAGGAATTATTGGAATTGGACAAGT

AlaSerLeuTrpAsnTrpPheSerIleSerLysTrpLeuTrpTyrIleArgIlePheIleGGGCAAGTTTGTGGAATTGGTTTAGCATATCAAAATGGCTGTGGTATATAAGAATATCA

ArgValArgGlnGlyTyrSerProLeuSerLeuGlnThrLeuLeuProThrProArgGlyATAGAGTTAGGCAGGGGATACTCACCTCTGTCGTTGCAGACCCTCCTCCCAACACCGAGGG

ProFroAspArgProGluGlyIleGluGluGluGlyGlyGluGlnGlyArgGlyArgSer GACCACCCGACAGGCCCGAAGGAATAGAAGAAGGAGGTGGAGAGCAAGGCAGAGGCAGAT 8000

IleArgLeuValAsnGlyPheSerAlaLeuIleTrpAspAspLeuArgAsnLeuCysLeuCAATTCGATTGGTGAACGGATTCTCAGCACTTATCTGGGACGACCTGAGGAACCTGTGCC

PheSerTyrHisArgLeuArgAspLeuLeuLeuIleAlaThrArgIleValGluLeuLeu
TCTTCAGTTACCACCGCTTGAGAGACTTACTCTTAATTGCAACGAGGATTGTGGAACTTC

GlyArgArgGlyTrpGluAlaLeuLysTyrLeuTrpAsnLeuLeuGlnTyrTrpGlyGln-TGGGACGCAGGGGGGGGAAGCCCTCAAATATCTGTGGAATCTCCTGCAATATTGGGGTC



GluLeuTysAsnSerAlaIleSerLeuLeuAsnThrThrAlaIleAlaValAlaGluCys AGGAACTGAAGAATAGTGCTATTAGCTTGCTTAATACCACAGCAATAGCAGTAGCTGAAT ThrAspArgValIleGluIleGlyGlnArgPheGlyArgAlaIleLeuEisIleProArg GCACAGATAGGGTTATAGAAATAGGACAAAGATTTGGTAGAGCTATTCTCCACATACCTA 8300 HeEGlyGlyLysTrpSerLys EW. ArglleArgGlnGlyPheGluArgAlaLeuLeu GAAGAATTAGACAGGGCTTCGAAAGGGCTTTGCTAFAACATGGGTGGCAAGTGGTCAAAA SerSerIleValGlyTrpProLysIleArgGluArgIleArgArgThrProProThrGlu AGTAGCATAGTAGGATGGCCTAAGATTAGGGAAAGAATAAGACGAACTCCCCCAACAGAA ThrGlyValGlyAlaValSerGlnAspAlaValSerGlnAspLeuAspLysCysGlyAla ACAGGAGTAGGAGCAGTATCTCAAGATGCAGTATCTCAAGATTTAGATAAATGTGGAGCA 8500 AlaAlaSerSerSerProAlaAlaAsnAsnAlaSerCysGluProProGluGluGluGlu GCCGCAAGCAGCAGTCCAGCAGCTAATAATGCTAGTTGTGAACCACCAGAAGAAGAAGAG GluValGlyPheProValArgProGlnValProLeuArgProMetThrTyrLysGlyAla GAGGTAGGCTTTCCAGTCCGTCCTCAGGTACCTTTAAGACCAATGACTTATAAAGGAGCT PhcAspLeuSerHisPheLeuLysGluLysGlyGlyLeuAspGlyLeuValTrpSerPro TTTGATCTCAGCCACTTTTTAAAAGAAAAGGGGGGACTGGATGGGTTAGTTTGGTCCCCA 8700 LysArgGlnGluIleLeuAspLeuTrpValTyrHisThrGlnGlyTyrPheProAspTrp AAAAGACAAGAAATCCTTGATCTGTGGGTCTACCACACAAGGCTACTTCCCTGATTGG GlnAsnTyrThrProGlyProGlyIleArgPheProLeuThrPheGlyTrpCysPheLys CAGAATTÁCACACCAGGGCCAGGGATTAGÁTTCCCACTGACCTTCGGÁTGGTGCTTTAÁG 0088 LeuValProHetSerProGluGluValGluGluAlaAsnGluGlyGluAsnAsnCysLeu TTAGTACCAATGAGTCCAGAGGAAGTAGAGGAGGCCAATGAAGGAGAGAACAACTGTCTG LeuRisProIleSerGlnEisGlyHetGluAspAlaGluArgGluValLeuLysTrpLys TTACACCCTATTAGCCAACATGGAATGGAGGACGCAGAAAGAGAAATGGAAG 8900 PheAspSerSerLeuAlaLeuArgHisArgAlaArgGluGlnEisProGluTyrTyrLys TTTGACAGCAGCCTAGCACTAAGACACAGAGCCAGAGAACAACATCCGGAGTACTACAAA AspCys GACTGCTGACACAGAAGTTGCTGACAGGGGACTTTCCGCTGGGGACTTTCCAGGGGAGGC GTAACTTGGGCGGGACCGGGGAGTGGCTAACCCTCAGATGCTGCATATAAGCAGCTGCTT $JJ \stackrel{\longrightarrow}{\longrightarrow} R$ CTTGTTAGACCAGGTCGAGCCCGGGAGCTCTCTGGCTAGC TTCGCCTGTACTCCGTCTCT AAGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCCTCAA

9200